Chalmers, James McNish
An analysis of results obtained on the Wechsler intelligence scale for children by mentally superior subjects.

Faculty of Education.

THE UNIVERSITY OF ALBERTA

AN ANALYSIS OF RESULTS OBTAINED ON THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN BY MENTALLY SUPERIOR SUBJECTS

A DISSERTATION SUBMITTED

TO THE COMMITTEE ON GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE DEGREE OF

MASTER OF EDUCATION

FACULTY OF EDUCATION

BY

JAMES McNISH CHALMERS

EDMONTON, ALBERTA

APRIL, 1953

UNIVERSITY OF ALBERTA

FACULTY OF EDUCATION

The undersigned hereby certify that they have read and do recommend to the Committee on Graduate Studies for acceptance, a thesis entitled "An Analysis of Results Obtained on the Wechsler Intelligence Scale for Children by Mentally Superior Subjects", submitted by James McNish Chalmers, B.Sc.,B.Ed., in partial fulfilment of the requirements for the degree of Master of Education.

Professor

Professor

Professor

Professor

ACKNOWLEDGMENTS

During the preparation of this thesis the writer received assistance and advice for which he wishes to express sincere appreciation. In particular, he is obligated to Dr. G. M. Dunlop, Chairman, Division of Educational Psychology, Faculty of Education, under whose direction the work was developed and completed. The helpful criticisms and suggestions of Dean H. E. Smith and Dr. R. E. Rees, Faculty of Education, and of Professor E. S. Keeping and Professor T. Fostvedt, Department of Mathematics, are gratefully acknowledged.

In addition, the writer was assisted immeasurably in the testing program through the willing co-operation of the following directors and principals of the Edmonton Public Schools: Mr. G. R. Conquest, Director of Secondary Education and Guidance, Mr. A. G. Bayly, Director of Elementary Education, Mr. J. G. Sinclair, Mr. A. O. Millions, Mr. E. M. Beckwith, and Mr. C. J. Allison.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

TABLE		PAGE

TABLE		PAGE

TABLE		PAGE

LIST OF FIGURES

SYNOPSIS

The main purpose of this study was to examine the results obtained on the Wechsler Intelligence Scale for Children by a group of mentally superior subjects. The test was administered to a number of Edmonton public school students who had previously obtained high scores on various other mental tests, and to a few of the brighter children attending school in the village of Carstairs, Alberta. Only those subjects who rated a WISC Full Scale I.Q. of at least 120 were included in the study.

Analyses included examination of the magnitude and significance of rank correlation coefficients of the eleven subtest scores with the Full Scale score, and of the Verbal and Performance subtest scores with the Verbal and Performance total scores respectively. The significance of the differences existing among the means of the eleven subtest scores was computed. Attention was given to the means of the Verbal and Performance Intelligence Quotients obtained by the group, while an ancillary study was undertaken to determine the nature of the distribution of discrepancies occurring between Verbal and Performance Intelligence Quotients. Raw scores obtained by the subjects were examined in an effort to determine which, if any, of the WISC subtests are of insufficient difficulty to measure adequately the intelligence of

superior children.

The investigation revealed that in most cases the rank correlation coefficients of the subtest scores with the Full Scale score were quite low, often failing to be significantly different from zero. Correlation coefficients of the Verbal subtests with total Verbal score, and of the Performance subtests with total Performance score were somewhat higher, but no coefficient approached the Pearson's r values published by Wechsler for his standardization sample.

There was a rather pronounced tendency for these superior children to obtain significantly higher weighted scores on the Verbal subtests than on the Performance subtests. As a result, the mean Verbal I.Q. of the group was found to be substantially greater than the mean Performance I.Q. The ancillary study dealing with the deviations of Verbal from Performance Intelligence Quotients tended to bear out this latter finding, as the distribution of discrepancy scores indicated that Verbal Intelligence Quotients were commonly greater than Performance Intelligence Quotients.

Raw scores obtained by many of the subjects point to the likelihood that several of the WISC subtests do not possess sufficiently high ceilings to measure fully the mental abilities of certain superior adolescents. In addition, it was felt that Wechsler's procedure of limiting

the magnitude of scaled scores to a maximum of twenty points tended to hinder accurate measurement of individual differences in cases where especially high raw scores were obtained.

CHAPTER I

INTRODUCTION

Interest in the measurement of human characteristics dates back to the early years of the nineteenth century. Groups working in the fields of physiology, anthropology, and psychology gradually became more conscious of individual differences and began to search for methods that would enable them to measure the extent and nature of these differences. Darwin, Galton, Wundt, Cattell, Wissler, and Kraeplin were outstanding among those who laid the foundations of this movement, and their testing procedures, although often crude and inaccurate, gave later scientists something on which to build.

Toward the end of the nineteenth century a French psychologist, Alfred Binet, began experimenting with tests which would measure such mental functions as memory, reasoning, and attention. From the time of his first test series in 1896, he insisted that mental tests should be designed to measure a variety of higher processes, and should not emphasize sensory discrimination and reaction.[1] His early emphasis upon tests of memory and imagery suggest that these mental functions were regarded as close simulators of intelligence.[2] Successively as his work progressed

he devised tests of attention, adaptation, judgment, direction, auto-criticism, comprehension, and invention.[3] In his tireless efforts to discover the nature of intelligence, Binet did not allow himself to be too greatly bound by his earlier theories, and he showed considerable willingness to cast aside any view that did not test out well. In the end, after trying vainly to disentangle the various intellective functions, he decided to test their combined functional capacity without any pretense of measuring the exact contribution of each to the total product.[4]

The school officials of Paris had been watching Binet's work with interest. In 1904 they asked him to assemble a battery of tests that would assist them in differentiating between the genuinely dull pupils and those who were capable of absorbing regular classroom instruction. With the assistance of Theodore Simon, Binet published his first scale of general intelligence in 1905. This battery consisted of a list of thirty tasks to be performed by the child, and was constructed with the primary object of affording a method by which mentally deficient students could be objectively and accurately selected. Further research led to the 1908 and 1911 revisions of the original scale, in which Binet introduced[5] and perfected the method of grouping the test items by age.[6]

The prestige of these intelligence scales increased

steadily as foreign psychologists came to realize their value. After some six years of research, Lewis Terman published in 1916 the Stanford Revision of the Binet Scale. Additional items in this revised battery supplemented Binet's original tests to the extent that it was now possible to measure adequately the intellectual capacities of normal and superior children of widely varying ages. The Stanford Revision had immediate popularity, and became for the next twenty years the yardstick by which other tests were judged. In 1937 Terman and Merrill introduced the Revised Stanford-Binet Scale, which is today the most widely used test of general intelligence on the market.

Like the earlier Binet Scales, the present scale is strongly weighted with verbal items, and has been found to be discriminatory in cases where there is evidence of poor schooling, linguistic handicap, or reading disability.[7] Because of these and other weaknesses discovered in the general design of the Revised Stanford-Binet Scale, David Wechsler, clinical psychologist at New York's Bellevue Hospital, constructed a test more appropriate to his needs. The Wechsler-Bellevue Scale, as it was called, provided some means of revealing more about the pattern of an individual's mental functioning. With verbal and performance abilities weighted equally, it was now possible to calculate a Verbal, a Performance, and a Full Scale I.Q. for

each subject. In 1946, seven years after the appearance of the first edition, Wechsler published Form II of the test. This new form is essentially the same in design as the original, but it has different test items, and thus affords a reliable means of retesting the individual.

The Adult Scales earned such favorable publicity that the author decided to construct a similar test for use with younger subjects. This new battery, the Wechsler Intelligence Scale for Children, is made up largely of items taken from Form II of the earlier scales. The main additions are new items at the easier end of each subtest to permit examination of children as young as five years of age. Wechsler contrasts the Bellevue Scales and the WISC as follows:

> Even though the materials overlap, the WISC is a distinct test from the Wechsler-Bellevue Scales, and is independently standardized. The directions for presenting and the standards for scoring and assignment of weights and bonuses are quite different.[6]

The WISC is composed of 12 subtests, of which 6 are classified as Verbal, and 6 as Performance. The Verbal Scale includes tests of Information, Comprehension, Arithmetic, Similarities, Vocabulary, and Digit Span, while the Performance subtests are known as Picture Completion, Picture Arrangement, Block Design, Object Assembly, Coding, and Mazes. Ordinarily only 5 Verbal and 5 Performance tests are administered to the subject, and the intelligence

quotients are calculated on this basis. Digit Span and Mazes are usually omitted because of their low correlations with the other tests of the scale. While it is permissible to give all tests, the total weighted scores obtained require prorating before the Intelligence Quotients may be computed from the tables.

According to Wechsler, general intelligence is not a unitary trait or ability, but is part of personality itself. It is something more than mental energy or the power to educe relations: it is the ability to utilize this mental energy or to exercise this power in contextual situations.[9] He argues that tests of verbal ability, abstract reasoning and the like, when used alone in a general intelligence examination, have been found to give an incomplete picture of the individual's capacity for effective adjustment and achievement. As a consequence, he has made a deliberate attempt in the construction of the WISC to take into account certain non-intellective factors of the personality such as persistence and desire to succeed. These factors, he contends, contribute to the total effective intelligence of the individual, and are objectively appraisable in the various performance tests.[10]

CHAPTER II

PURPOSE OF THE INVESTIGATION

There is abundant evidence to indicate that psychological and educational authorities are neglecting the child with superior intelligence. The amount of research conducted with normal and subnormal children has far outweighed the studies directed toward the understanding of the gifted.

> The American Association for Gifted Children is conducting a wide range of activities designed to meet some of the outstanding needs of this group. Among the most urgent needs are: a more widespread understanding of the nature of gifted children and youth, more efficient teachers, improved working relationships between parents and teachers, more varied and more stimulating curricula in our school system, and more research on the gifted.[11]

A glance at indices of newspapers, periodicals, and even professional journals reveals the fact that references to the gifted child are very few.

> A graduate student at Northwestern University recently consulted the Educational Index to discover the number of articles on gifted children listed for the years 1929-48. She found that the peak of interest in gifted children as reflected in the number of articles published was during the years 1939-42. Interest lagged throughout the period of World War II and a rather consistent average was reached in the years directly following.[12]

The space allotted to the gifted in educational and psychological publications offers another indication of the small amount of attention that has been given to this group.

For example, in the volume on Special Education, "The Handicapped and the Gifted", more than 515 pages were devoted to the handicapped, while only 13 pages were given over to the gifted pupils.[13]

In the provision of special classes for abnormal groups, the trend is again very much in the same direction.

> Compared with the number of classes organized for the subnormal child, special class work for the superior is woefully lacking. Healihy reported only 2,842 children attending special classes for the superior in 1938, as contrasted to about 126,000 subnormals for whom special educational provisions had been made.[14]

> The table of statistics for 1947-48 furnished by the United States Office of Education ... shows that fewer than 21,000 gifted children are enrolled in special classes in the elementary and the secondary schools.[15]

The history of intelligence testing reflects once again the apparent tendency to overlook the superior individual, and to concentrate chiefly on the abilities of the normal and subnormal groups.

> Binet held that the most important application of the tests of intelligence was to persons of inferior intelligence. He did not recognize explicitly the great value of testing for superior intelligence and for outlining for super-normal children an education according to their abilities ...[16]

> For Wechsler, as for Binet, the principal interest was in studying the average and below-average groups; no particular effort was made to measure precisely the higher levels of adult mental ability.[17]

One of the chief criticisms of the Wechsler-Bellevue

Scale is that it is not sufficiently difficult for measurement of superior subjects.[18] Has this weakness been remedied in the newer children's scale by the addition of items of appropriate complexity at the upper end of each test? The instructor of mathematics may be inclined at first glance to doubt whether the Arithmetic subtest is composed of problems that will tax the mental agility of more capable adolescents. A number of the writer's professional colleagues felt, upon examining the WISC for the first time, that the most difficult items on the Similarities, Comprehension, and Picture Completion subtests were not of a quality that would challenge the intelligence of certain exceptional children of their acquaintance.

It was disclosed about eighteen months ago that a number of superior children tested at the University of Alberta's Psychological Clinic had obtained noticeably deviant scaled scores on the various WISC subtests. One testee, for example, had secured the maximum scaled score of 20 points on the Similarities subtest but had fallen to 8 on Coding, while another had scored 19 on Comprehension and only 10 on the Block Design and Coding subtests. Were widely varying scores such as these to be regarded as the normal achievement of mentally alert children? The more recent psychological journals were consulted at that time to determine what findings had been made regarding this

phenomenon. The search revealed that while a number of studies of normal and feeble-minded groups involving the Wechsler-Bellevue Scales and the WISC had been undertaken, none was on record that was primarily concerned with the attainments of superior subjects on these scales. This disclosure aroused sufficient interest and curiosity to warrant a fuller inquiry into the matter.

The present investigation was undertaken, therefore, in an effort to assemble information concerning the achievements of mentally superior children on the WISC, and to compare the results with those obtained by Wechsler from his standardization sample. It was conducted with the desire to fill a gap in the body of research findings on the new scale, and perhaps make some contribution to the understanding and interpretation of the WISC at the upper levels of intelligence.

CHAPTER III

RELATED STUDIES

The dearth of published articles dealing with the achievements of superior children on the WISC is not surprising when one realizes that this test has been on the market for a comparatively short period of time. Again, as mentioned in a previous chapter, there is evidence to indicate that the feeble-minded child has been subject to closer scrutiny than has the gifted child. This situation is due perhaps to the assumption on the part of educators and test administrators that the individual who possesses superior intelligence requires less assistance and attention than the mentally inferior person. It is felt by the writer that the combination of these two factors explains in part the relative paucity of published findings related to this investigation.

I. RELEVANT CONCLUSIONS DRAWN BY WECHSLER

Upon analyzing the test profiles obtained in the Wechsler-Bellevue standardization study, Wechsler was able to draw a number of conclusions regarding the achievement of deviant groups. He found, for instance, that subjects with superior intelligence generally did better on the

verbal subtests, while mentally inferior individuals tended to obtain their highest scores on the performance section of the examination. In addition, he confirmed the findings of previous studies which indicated that Jewish children did better on verbal tests, and Italian children on performance tests.[19]

II. INTELLIGENCE QUOTIENTS OF SUPERIOR SUBJECTS IN THE WISC STANDARDIZATION SAMPLE

The superiority of Verbal I.Q. to Performance I.Q. among gifted subjects was substantiated in the process of standardizing the WISC. Of the 2200 cases included in the sample, 2.1% were found to possess a Verbal I.Q. of 130 points or more, whereas the corresponding percentage on the Performance Scale was only 1.1%. The number of children who tested over 120 I.Q. on the three scales is indicated in Table I.

It may be noted that the Verbal Scale cumulative frequencies exceed those of the Performance Scale at all of the seven levels of superiority listed in Table I. One would gather from this that Verbal Intelligence Quotients tend to be greater in magnitude than Performance Intelligence Quotients among superior children. This tendency appears to be especially marked where the Intelligence Quotients exceed 130.

TABLE I. DISTRIBUTION OF VERBAL, PERFORMANCE, AND FULL SCALE INTELLIGENCE QUOTIENTS AT VARIOUS LEVELS OF SUPERIORITY IN THE WISC STANDARDIZATION SAMPLE

WISC I.Q.	NUMBER OF SUBJECTS EXCEEDING INDICATED I.Q.		
	VERBAL SCALE	PERFORMANCE SCALE	FULL SCALE
150+	1	0	0
145+	4	1	2
140+	11	3	4
135+	21	11	14
130+	46	25	32
125+	99	84	83
120+	199	195	179

(The above information was taken from Tables XI, XII and XIII, Seashore, and others)[20]

III. DISCREPANCY SCORES OF SUBJECTS CONSTITUTING THE WISC STANDARDIZATION SAMPLE

Seashore has carried out considerable research on the WISC's standardization sample, and recently analyzed the size and direction of discrepancies between the Verbal and Performance Intelligence Quotients of the 2200 subjects tested.[21] He defined the term 'discrepancy score' as the numerical difference between a child's two Intelligence Quotients, with Performance I.Q. always being subtracted from Verbal I.Q. Both positive and negative discrepancy scores are thus obtainable. Seashore found that the means of the discrepancy scores at each age level from 5 - 15 years were all essentially zero, with the standard devia-

tions of all distributions being about 12 or 13 I.Q. units. This indicated that approximately one-third of the discrepancies between Verbal and Performance Intelligence Quotients were greater than 12 - 13 points of I.Q., while in about two-thirds of the cases the differences ranged from zero to 12 - 13 points. Some of the 2200 discrepancy scores were quite large, ranging to extreme values of approximately -35 and +35.

There appeared to be a slight tendency for urban children to have positive discrepancy scores, and for rural subjects to possess relatively higher Performance Intelligence Quotients than Verbal Intelligence Quotients. It was further found that children with professional and semi-professional backgrounds tended to form the most verbal-minded group, although once again the trend was not too marked. The 55 subjects in the sample's subnormal group obtained a mean discrepancy score of -2. This deviation from zero, however, was not considered to be great enough to be of clinical importance. No reference was made to the nature of the discrepancies between Verbal and Performance Intelligence Quotients among the highly intelligent children in the sample.

IV. DISCREPANCY SCORES OF YOUNGER SUBJECTS IN THE WECHSLER-BELLEVUE STANDARDIZATION SAMPLE

Wechsler was well aware that fairly large discrepancies

existed between the Verbal and Performance Intelligence Quotients of the subjects composing the Wechsler-Bellevue standardization sample. The observed differences for the 520 subjects in the 10 - 16 age group were analyzed from the standpoint of absolute magnitude and the median discrepancy score regardless of sign was found to be 9.1 points of I.Q.[22] Forty-one subjects in this group possessed Full Scale Intelligence Quotients which exceeded 120. The median absolute difference between these 41 pairs of Verbal and Performance Intelligence Quotients was found to be 14.9 points of I.Q.[23] Table II indicates the percent of cases in which Verbal and Performance Intelligence Quotients either outweighed each other or were identical in magnitude.

TABLE II. PERCENT OF CASES IN THE 10 - 16 AGE GROUP OF THE BELLEVUE STANDARDIZATION SAMPLE WITH HIGHER VERBAL AND HIGHER PERFORMANCE INTELLIGENCE QUOTIENTS

I.Q. Category	Number of Cases	Performance I.Q. Higher	Verbal I.Q. Higher	No Difference
79 and below	40	67.5%	25.0%	7.5%
91 - 110	287	52.3%	43.9%	3.8%
120 and over	41	12.2%	82.9%	4.9%

(The above information was taken from Tables XXIV and XXV, Wechsler.)[24]

Note that the total number of cases considered in Table II is 368, not 520. This difference in total fre-

quencies is due to the fact that the I.Q. categories 80 - 90 and 111 - 119 are not included in the table.

An inspection of the percentage frequencies within the three I.Q. categories in Table II would lead one to conclude that the proportion of positive Bellevue discrepancy scores varies directly as the Full Scale I.Q. level of a homogeneous group of subjects.

V. STUDIES INVOLVING THE WISC AND MENTALLY DEFICIENT CHILDREN

A correlational analysis of the scores of 70 subnormal subjects on the Revised Stanford-Binet Scale and the WISC was reported by Stacey and Levin.[25] The children were divided into two groups, one composed of 44 morons and the other of 26 borderline cases. In the moron group a coefficient of .21 was found when Verbal subtest totals were correlated with Performance totals, while the corresponding coefficient for the borderline group was -.31.

In Taylor's study of 74 feeble-minded children conducted at the same time as the present investigation, the following coefficients of rank correlation were obtained:[26]

Information	vx.	9 other subtests..	.29
Comprehension	vs.	9	.15
Arithmetic	vs.	9	.40
Similarities	vs.	9	.36
Vocabulary	vs.	9	.29
Digit Span	vs.	10	.17

Picture Completion	vs. 9	.21
Picture Arrangement	vs. 9	.31
Block Design	vs. 9	.33
Object Assembly	vs. 9	.14
Coding	vs. 9	.29
Total Verbal vs. Total Performance		.21

The correlational results obtained in these two investigations of feeble-minded subjects have been included in the present list of related studies merely to suggest that there may be a lesser degree of correlation between the various combinations of WISC scores in intellectually homogeneous groups than Wechsler has indicated for his standardization sample.

CHAPTER IV

POPULAR CLASSIFICATIONS OF SUPERIOR INTELLIGENCE

Those persons and groups who have taken upon themselves the task of constructing intelligence scales have commonly agreed that the concept of intelligence levels has proven very convenient in the classification of mental ability. Unfortunately, however, little harmony has existed either in nomenclature or definition of the various levels of intelligence.

Of current test-makers, none is more aware of these inconsistencies than Wechsler, who disapproves of the arbitrary bases on which Terman, Kuhlmann, and others have classified their Intelligence Quotients.[27] He questions the wide acceptance of the Terman classificatory scheme, not only for ratings obtained on the Stanford-Binet, but for Intelligence Quotients derived from any number of other tests. Furthermore, he condemns as unsound the practice of reporting equivalent Binet I.Q. ratings for nearly every intelligence scale currently in use, as this procedure entails the blind assumption that these tests are measuring the same factor or factors as the Binet.[28]

Wechsler wished to construct a classificatory scheme in which the definition of intelligence levels was in terms

of statistical frequencies. Realizing the popularity of Terman's system, he attempted to retain as much of it as possible while at the same time introducing the statistical quantity 'Probable Error' into the determination of class limits.[29] He argued that an average person was generally considered to be one who fell within the middle 50 percent of the group, a range which was defined on the normal curve baseline by the interval -1 P.E. to +1 P.E. Concluding that an I.Q. of 90 was interpreted in the case of most intelligence scales as the lowest limit of average or normal intelligence, he equated the P.E. value of -1 with 90.[30] This meant that if the mean intelligence quotient were set at the traditionally accepted value of 100, a distance of 1 P.E. on the normal curve baseline would then correspond to 10 points of I.Q.

As a means of ascertaining the validity of the proposed scheme, Wechsler set the lower limit of borderline deficiency at -3 P.E., a point which corresponded to the value adopted by Terman, viz. 70 I.Q. A mental defective would then be a person falling at a distance greater than 3 P.E.'s from the mean. An ordinate erected at -3 P.E. was found to cut off the lower 2.2% of the area under a normal curve, indicating that this percentage of a normal population would be classified as mentally deficient under the proposed scheme. With

the knowledge that various estimates of the probable incidence of mental deficiency in America gave a mean figure approximating 3% of the total population, Wechsler was apparently satisfied that this classificatory procedure was reasonably valid.[31] As a result, WISC I.Q. tables have been constructed in such a manner as to provide, for a large randomly-selected group, a normal distribution of quotients having a mean of 100 and a standard deviation of 15, which is to say a P.E. of 10.[32]

Wechsler's recently revised classification of intelligence is presented in Table III. For the purpose of comparison, Terman's classification is shown in Table IV.

TABLE III. CLASSIFICATION OF INTELLIGENCE ACCORDING TO WECHSLER

Classification	Limits in Terms of P.E.	I.Q. Range	Per Cent Included
Very Superior	+3 P.E. and over	130 and above	2.2
Superior	+2 P.E. to +3 P.E.	120 - 129	6.7
Bright Normal	+1 P.E. to +2 P.E.	110 - 119	16.1
Average	-1 P.E. to +1 P.E.	90 - 109	50.0
Dull Normal	-2 P.E. to -1 P.E.	80 - 89	16.1
Borderline	-3 P.E. to -2 P.E.	70 - 79	6.7
Mental Defective	-3 P.E. and below	69 and below	2.2

(The above information was taken partly from Table IX of the WISC Manual[33], and partly from Table III, Wechsler.[34])

TABLE IV. CLASSIFICATION OF INTELLIGENCE ACCORDING TO TERMAN

Classification	I.Q. Range
Genius or Near Genius	140 and above
Very Superior Intelligence	120 - 140
Superior Intelligence	110 - 120
Normal or Average Intelligence	90 - 110
Dullness	80 - 90
Borderline Deficiency	70 - 80
Definite Feeble-mindedness	Below 70

(Taken from Terman.[35])

Consideration of Tables III and IV indicates that Terman and Wechsler have agreed quite closely in their classifications of intelligence at the subnormal and normal levels. A marked difference is apparent, however, in the two methods of classifying the mentally superior. At least one writer has drawn attention to certain possible misunderstandings that may arise from these dissimilar classifications of above average intelligence.[36]

Recent studies have indicated that there is probably even greater discrepancy at the upper levels than Tables III and IV would suggest. The results obtained in one investigation disclosed that the mean score of subjects who rated more than 130 I.Q. on the Revised Stanford-Binet was nearly 20 points in excess of that obtained by the same group on the WISC Full Scale.[37] In a somewhat similar study, a trio of investigators constructed a table

of equivalent Stanford-Binet and WISC Intelligence Quotients. This table indicates that WISC Full Scale Intelligence Quotients of 120 and 130 are respectively equivalent to Revised Stanford-Binet Intelligence Quotients of 128 and 140.[38]

Whatever the true discrepancies may be between the I.Q. scores on the two scales at the upper levels of intelligence, there is apparently a very good chance that a Bright Normal child of WISC I.Q. 115 may obtain a grading of Very Superior when tested and classified on the Revised Stanford-Binet Scale. The only conclusion to be drawn is that these classificatory schemes differ markedly at the higher levels of intelligence.

CHAPTER V

RESEARCH PROCEDURES

I. THE SELECTION OF SUBJECTS

Since the purpose of this investigation was to examine the nature of the results obtained on the WISC by a group of mentally superior children, the decision was made to abide by Wechsler's classificatory scheme, and to accept 120 I.Q. on the Full Scale as the lower limit of the superior group. The upper limit was left undefined inasmuch as both superior and very superior children were under consideration. Since the WISC is an individual test of intelligence that requires approximately eighty minutes for proper administration and scoring, it was deemed necessary at the outset to adopt some method of screening those subjects who were not likely to obtain quotients of sufficient magnitude to be of use in the investigation.

Two group intelligence tests, namely the Detroit Beginning First-Grade and the Laycock Mental Ability scales, are administered annually to pupils entering grades I and V respectively in the Edmonton Public School system. The writer had access to a number of these scores, and found them of considerable value in eliminating children who possessed less than superior intelligence. Upon careful

examination of these group scores and after conferring with many of the room teachers involved, it was decided to test only those students who had obtained I.Q. ratings of at least 125 on either or both of these scales. In a few outside cases, however, the WISC was administered where it was felt that the Detroit or Laycock score had underestimated the true mental ability of the child. The screening measures proved quite successful on the whole, with only 9 of 66 testees failing to obtain the minimum standard of 120 I.Q. on the Full Scale.

It was felt that this method of elimination had the effect of somewhat diminishing the randomness of the sample chosen, since many subjects who might have scored just over 120 I.Q. were not included in the survey. However, the only means of securing a truly random sample would have necessitated testing a very large number of children and then discarding all cases that fell below 120 I.Q. This procedure was not attempted because of the exhorbitant consumption of time entailed.

II. DISTRIBUTION OF INTELLIGENCE QUOTIENTS IN THE SAMPLE

In one of his publications Seashore has indicated the distribution of 179 Superior I.Q. scores obtained from the WISC's standardization sample.[39] The frequencies of these

TABLE V. DISTRIBUTIONS OF SUPERIOR I.Q. SCORES OBTAINED IN THE WISC STANDARDIZATION SAMPLE AND IN THE CURRENT STUDY

WISC Full Scale Intelligence Quotients	Frequencies	
	Standardization Sample	Current Sample
150-154	0	0
145-149	2	2
140-144	2	2
135-139	10	5
130-134	18	9
125-129	51	24
120-124	96	15
Total N:	179	57

TABLE VI. COMPARISON OF PERCENTAGE FREQUENCY DISTRIBUTIONS OF WISC SUPERIOR I.Q. SCORES OBTAINED FROM TWO SAMPLES WITH THEORETICAL VALUES DERIVED FROM A P.E. TABLE

I.Q. Range	Percentage Frequencies		
	Based on P.E. Values	Standardization Sample	Current Sample
155 and above	0.1%	0.0%	0.0%
150-154	0.3%	0.0%	0.0%
145-149	0.9%	1.1%	3.5%
140-144	2.6%	1.1%	3.5%
135-139	6.3%	5.6%	8.8%
130-134	14.0%	10.1%	15.8%
125-129	27.5%	28.5%	42.1%
120-124	48.3%	53.6%	26.3%
Totals:	100.0%	100.0%	100.0%

scores are grouped in intervals of 5 I.Q. points and compared in Table V with the corresponding frequencies realized in the present study.

Percentage frequencies were calculated for the two distributions of superior I.Q. scores, and these are compared in Table VI with expected values derived from a table of P.E.

III. COMPUTING THE THEORETICAL DISTRIBUTION OF SUPERIOR INTELLIGENCE QUOTIENTS

As mentioned in a previous chapter, in the construction of his classificatory scheme, Wechsler made the assumption that intelligence was a trait normally distributed about the traditionally accepted mean of 100 I.Q. He then chose to equate 10 points of I.Q. to a distance of 1 P.E. along the baseline of a normal probability curve. Persons of average intelligence were considered to be those whose Intelligence Quotients fell within the middle 50% of a distribution of scores obtained from a large, randomly selected sample of subjects, a range defined on the normal curve baseline by the distance -1 P.E. to +1 P.E. Successive points were then laid off on the baseline in units of P.E., following which a categorical title was assigned to each of these intervening distances.

Wechsler's recently revised classificatory scheme is

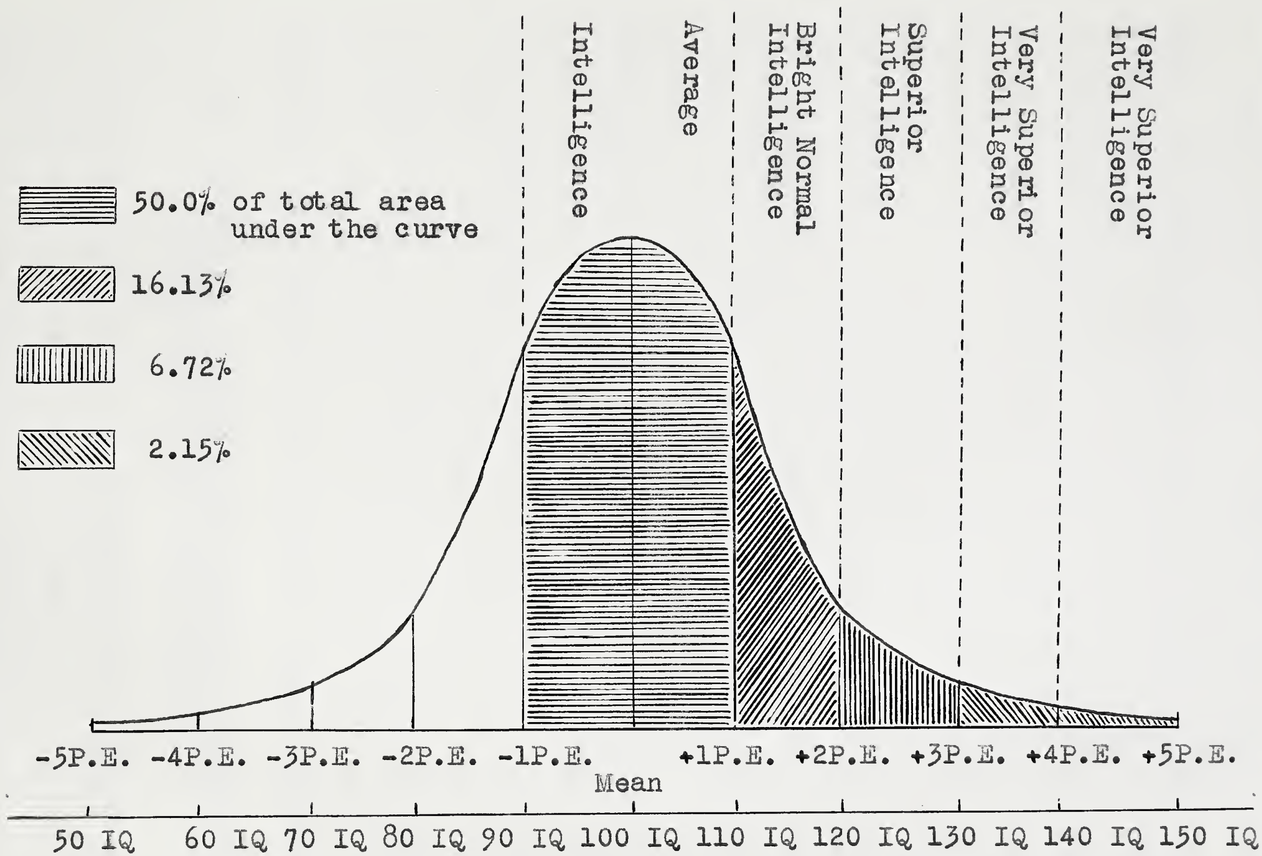

Figure 1. Classification of Average and Above-average Intelligence According to Wechsler

indicated diagrammatically in Figure 1.

From a table of P.E., it may be seen that 8.87% of cases in a normal distribution lie above +2 P.E., while 4.28% fall within the distance +2 P.E. to +2.5 P.E.[40] Accordingly, 428/887 or 48.3% of the cases lying above +2 P.E. are located within the range +2 P.E. to +2.5 P.E. Translated into Wechsler's classificatory scheme of intelligence, this suggests that 48.3% of a large, randomly selected group of Superior and Very Superior subjects should fall within the interval 120-124.9 I.Q., while 51.7% of these subjects should attain Intelligence Quotients of 125 or more. By means of similar arguments it may be shown that 27.5% of all Superior and Very Superior children should score between 125 and 129.9 I.Q. on the WISC Full Scale, 14.0% should obtain from 130 to 134.9 I.Q., 6.3% from 135 to 139.9 I.Q., 2.6% from 140 to 144.9 I.Q., 0.9% from 145 to 149.9 I.Q., 0.3% from 150 to 154.9 I.Q., and 0.1% from 155 I.Q. up. These are the percentage frequencies that appear in the 'Based on P.E. Values' column of Table VI.

IV. CRITICISM OF THE DISTRIBUTION OF INTELLIGENCE QUOTIENTS IN THE CURRENT SAMPLE

An examination of the percentage frequencies appearing in the three columns of Table VI will bring to light certain flaws in the distributions of Intelligence Quotients in both

Wechsler's and the current samples. The standardization sample had apparently too few cases at the 130-134 I.Q. level, and too many within the 120-124 I.Q. interval. The current sample on the other hand is substantially too large at the 125-129 I.Q. level, while a considerable deficiency is evident in the number of cases within the interval 120-124 I.Q.

V. A DESCRIPTION OF THE SUBJECTS CONSTITUTING THE CURRENT SAMPLE

The mean age of the 58 subjects considered in this investigation was 12 years 4 months. The distribution of ages is shown in Table VII.

TABLE VII. DISTRIBUTION OF AGES OF THE 57 MENTALLY SUPERIOR SUBJECTS TESTED ON THE WISC

Age in Years and Months	Sex of Subject		Total Frequency
	Male	Female	
7-0 to 7-11	1	0	1
8-0 to 8-11	1	0	1
9-0 to 9-11	2	4	6
10-0 to 10-11	0	1	1
11-0 to 11-11	6	5	11
12-0 to 12-11	7	6	13
13-0 to 13-11	8	8	16
14-0 to 14-11	5	2	7
15-0 to 15-11	0	1	1
Totals:	30	27	57

The 57 subjects were drawn from eight Edmonton schools

and from the Carstairs Public School. Table VIII shows the distribution of subjects by school and grade.

TABLE VIII. SCHOOLS AND GRADES FROM WHICH THE 57 SUBJECTS OF SUPERIOR INTELLIGENCE WERE DRAWN

School	GRADE									Totals
	II	III	IV	V	VI	VII	VIII	IX	X	
McKay Ave.			5			6	3	1		15
Oliver						4	7	2		13
Spruce Ave.							8	3		11
University	1		1	1		2	1	1	1	8
McDougall					2		2			4
Garneau		1								1
Rutherford						1				1
Mt. Carmel								1		1
Carstairs					3					3
Totals	1	1	6	1	5	13	21	8	1	57

VI. CALCULATION OF MEAN SCORES AND STANDARD DEVIATIONS

Means of subtest, total Verbal, total Performance, and Full Scale weighted scores, and of Verbal, Performance, and Full Scale Intelligence Quotients were determined by application of the formula[41]

$$M = A.M. + \left(\frac{\Sigma fx'}{N}\right) i$$

In all cases N was 57, with the exception of the Digit Span subtest, in which instance N was 52. This subtest was administered in addition to the 10 regular subtests to 52 of the 57 subjects merely for the purpose of obtaining informa-

tion regarding its adequacy at the upper levels of intelligence. Digit Span weighted scores played no part in the determination of either the Verbal or the Full Scale Intelligence Quotients.

Standard deviations of the various distributions of subtest weighted scores and of Intelligence Quotients were computed by the so-called short method. This involved the grouping of scores into frequency distributions, the calculation of deviations from assumed means, and the application of formula[42]

$$S.D. = i\sqrt{\frac{\sum(fx')^2}{N} - c^2}$$

VII. CALCULATION OF THE COEFFICIENTS OF CORRELATION

According to Wechsler, only 8.9% of a large, randomly selected population may be expected to possess Superior or Very Superior intelligence.[43] With over 91% of cases excluded at the outset of this investigation, the assumption could not be made that this sample represented a normally distributed parent population. This realization rendered inappropriate the use of the product-moment method of correlating subtest scores with Verbal, Performance, and Full Scale scores. The Spearman and the Kendall methods of rank correlation were considered, inasmuch as neither stipulated the assumption of normality in the parent popu-

lation. Kendall's method appeared to cover the treatment of tied ranks more fully than Spearman's, especially from the standpoint of testing the significance of coefficients. As a consequence, the Kendall method of rank correlation was applied throughout.

VIII. KENDALL'S METHOD OF RANK CORRELATION

When a number of individuals are arranged in order according to some quality which they all possess to a varying degree, they are said to be ranked. The arrangement as a whole is called a ranking in which each member has a rank.[44]

Of value in this investigation was the fact that Kendall has considered rank correlations in their own right, irrespective of the nature of the parent population.[45] His coefficient of rank correlation is denoted by the Greek letter τ to distinguish it from Spearman's ρ, which is derived by an altogether different method, and from Pearson's product-moment coefficient 'r'. Kendall has organized the mathematical procedures underlying the determination of 'tau' in such a manner that the coefficient increases in magnitude from -1 to +1 as agreement between the ranks increases. Following convention, the magnitude of tau is -1 when the two quantities being considered

are in perfect disagreement, and +1 where perfect agreement exists between the two rankings. For intermediate degrees of correspondence, tau's magnitude falls between these limiting values.[46]

Kendall's rank correlation coefficient is defined by [47]

$$\mathrm{tau}_a = \frac{S}{\frac{1}{2}n(n-1)}$$

where S is a score whose magnitude is determined by the amount of agreement between 2 rankings of n numbers. Where tied ranks exist, an alternative form of the coefficient is given by[48]

$$\mathrm{tau}_b = \frac{S}{\sqrt{\frac{1}{2}n(n-1) - T}\,\sqrt{\frac{1}{2}n(n-1) - U}}$$

in which case T is the correction for ties in one ranking, and U for ties in the other.

Since numerous tied ranks existed in all subtest and total scores obtained in this investigation, tau_b was determined in each of the 23 instances where the rank correlation coefficient was desired. For the convenience of the reader, a sample calculation of tau_b is included in Appendix C.

IX. CALCULATION OF THE SIGNIFICANCE OF THE TAU'S

Kendall has determined tests of significance of the two coefficients tau_a and tau_b, or rather of the values of S from which tau_a and tau_b are derived.

> A test of tau is equivalent to a test of the corresponding value of S, one being a multiple of the other, and we shall find it arithmetically more convenient to deal with S.[49]

With ties involved in neither ranking, he has defined the variance of S (i.e. σ_s^2) by[50]

$$\text{Var } S = 1/18\ n(n-1)(2n+5)$$

where n is the number of ranked scores. If there are ties of extent t in one ranking and u in the other, then the variance of the distribution obtained by correlating one ranking with all n possible arrangements of the other is given by[51]

$$\text{Var } S = 1/18\left[n(n-1)(2n+5) - \Sigma_t t(t-1)(2t+5) - \Sigma_u u(u-1)(2u+5)\right] + \frac{1}{9n(n-1)(n-2)}\left[\Sigma_t t(t-1)(t-2)\right]\left[\Sigma_u u(u-1)(u-2)\right] + \frac{1}{2n(n-1)}\left[\Sigma_t t(t-1)\right]\left[\Sigma_u u(u-1)\right]$$

σ_s is found by taking the square root of the calculated Var S. The probability that a given multiple of σ_s will be attained or exceeded is then read from a table prepared by Kendall for the purpose.[52]

Calculation of the significance of one of the tau's obtained in this study has been carried out in Appendix C.

X. ELIMINATION OF THE SPURIOUS ELEMENT PRESENT IN PART-WHOLE CORRELATION

Spurious correlation arises when a total score is correlated with a subscore which is a part of the total score. When calculating the correlation coefficients of WISC subscores with Verbal, Performance, and Full Scale scores, Wechsler eliminated spurious results by employing McNemar's corrective formula.[53] This formula, however, was not deemed applicable in the present survey, since tau's were being considered rather than r's.[54] An alternate procedure for eliminating spuriousness was adopted by correlating each subtest weighted score with the total weighted score minus the subtest involved. Thus Verbal and Performance subtests were correlated with the total of the four other Verbal and Performance subtests respectively, while all subtests were correlated with the total of the nine other subtests. Digit Span scores, not being integral parts of total Verbal nor Full Scale scores, were correlated directly with the totals of the five Verbal and of the ten regular subtest scores, with no correction being required for spuriousness.

This method of eliminating spuriousness, incidentally, was the one employed by Wechsler in the calculation of correlations for the Bellevue Scales.[55] It is probable that McNemar had not yet developed his corrective formula at the time.

XI. RELATIONSHIPS EXISTING AMONG KENDALL'S, SPEARMAN'S, AND PEARSON'S COEFFICIENTS OF CORRELATION

Since all correlation coefficients in this study have of necessity been computed by Kendall's method of rank correlation, a problem arises as to the comparison of the tau values with the r's commonly obtained in psychological investigations of this type. It may be stated at the outset that a correlation coefficient, whether it be r, tau, or rho, is merely a measure of the correspondence between two quantities, and as such, is affected by sampling fluctuation. The value of each coefficient is -1 where the two quantities are in perfect disagreement, and +1 in cases where there is perfect agreement. For intermediate degrees of correspondence, the three coefficients lie between these limiting values. A number of studies have been conducted in an effort to determine the relationship that might exist between any two of the coefficients.

Pearson has shown that if the universe is normal, the grade correlation ρ_g and the ordinary correlation r are related by the equation[56]

$$r = 2 \sin\left(\frac{\pi \rho_g}{6}\right)$$

Since ranks and grades are connected by a simple relation, it follows that the correlation between ranks is the same as the correlation between grades. Yule and Kendall state,

however, that since ranking often involves ties, and is in practice applied to comparatively small samples from universes of doubtful normality, the use of the above formula should be made with the utmost reserve.[57] They advise, furthermore, that it would probably be better to avoid the formula altogether, and to rely on the rank correlation coefficient.[58]

For the purpose of measuring correlation between ranks, tau appears to compare favorably with rho.[59] Different as they might seem to be from consideration of their methods of calculation, they are frequently found in practice to give numerical values which are remarkably close, even for low values of n.[60] By an empirical argument, Kendall and associates showed it to be likely that the product-moment correlation $r_{\tau\rho}$ between tau and rho is

$$r_{\tau\rho} = \frac{2(n+1)}{\sqrt{2n(2n+5)}}$$

for all values of the sample size n.[61]

To the knowledge of the writer, no procedure has as yet been developed whereby a tau calculated for a sample from a non-normal parent population may be translated directly into an r which assumes sampling from a normal or near-normal universe. There is reason to believe, however, that the degree of correspondence between r's and tau's of from near-zero to medium magnitude is rather high.

XII. COMPARISON OF THE MAGNITUDES OF TWO COEFFICIENTS COMPUTED BY DIFFERENT METHODS

The view is expressed in some mathematical circles that investigators have often employed the product-moment method of correlation in instances where the requirements for normality in the universe were not assured. It is likely that while research workers are generally aware of the availability of correlation methods other than Pearson's, some may consciously shun the application of rank correlation procedures in favor of the more widely accepted product-moment approach, regardless of the stipulation involving normality.

Whatever the actual situation may be, the writer could locate no related studies presenting correlational findings in terms of tau or rho. It was felt that at least one product-moment correlation should be computed in the present study merely to provide a basis for comparison with the corresponding tau value obtained. Total Verbal vs. total Performance scores had produced a tau of only -0.15, not significant at the .10 level of confidence. Pearson's method was applied on the identical data, resulting in a product-moment coefficient of -0.16. Once again the coefficient was not significantly different from zero. This finding tended to add weight to the previously adopted conclusion

that no correlation existed between total Verbal and total Performance scores for the current sample of superior subjects.

The computation of the product-moment coefficient is shown in Appendix C.

XIII. TESTING THE SIGNIFICANCE OF MEAN DIFFERENCES

The 13 distributions of subtest, total Verbal, and total Performance scores could not justifiably be regarded as normal in the parent populations, nor could it be assumed that these distributions possessed a common variance. Consequently, a problem arose as to the choice of the most appropriate procedure to be adopted in testing the significance of the differences between various pairs of sample means.

The mean scores being compared were not independent or uncorrelated, inasmuch as they were obtained from a single group of subjects rather than from different groups.[62] Furthermore, none of Garrett's formulas for computing the standard error of the difference between correlated means applied in a situation where functionally equivalent tests rather than a single test were administered to the same group.[63]

The procedure of comparison in pairs advanced by Brookes and Dick[64] was recommended by members of the De-

partment of Mathematics, University of Alberta, as an appropriate test of the significance of mean differences established in the present study. This test involved pairing scores obtained by the 57 subjects on any two subtests, finding the difference between each pair of scores, and computing the deviations of the 57 individual differences from the corresponding mean difference. The standard errors of the mean differences and subsequently the critical ratios of these differences to their standard errors were then calculated.

The Brookes and Dick formula employed in the computation of the various critical ratios or t's is stated as follows:[65]

$$t = \frac{\bar{d}}{\sqrt{\frac{\sum(d-\bar{d})^2}{n(n-1)}}}$$

where $\bar{d}$ is the observed difference between two means, and $\sqrt{\frac{\sum(d-\bar{d})^2}{n(n-1)}}$ the standard error of $\bar{d}$.

The significance of the difference between Comprehension and Vocabulary mean scores is determined in Table XXXVI, Appendix C.

XIV. COMPARISON OF DISCREPANCY SCORES WITH THOSE OBTAINED FROM THE STANDARDIZATION SAMPLE

Wechsler's Verbal, Performance, and Full Scale I.Q.

tables were constructed to render a mean I.Q. of 100 and a standard deviation of 15 at each age level.[66] Thus for any randomly selected group the mean discrepancy between Verbal and Performance Intelligence Quotients should be near zero. Upon analysis of Wechsler's standardization sample, Seashore found that at all age levels (V - P) discrepancy scores appeared to be normally distributed about means that approximated zero.[67]

The discrepancy scores of the 57 superior children tested in the present study were calculated, grouped in intervals of 9, and compared with those obtained by Seashore from the standardization sample, and by Taylor from a sample of 74 mentally deficient children. In order to simplify the comparison of these three distributions of discrepancy scores, a percentage frequency polygon was constructed for each distribution. These frequency polygons erected on a single set of axes constitute Figure 3, which may be found in Chapter VIII.

CHAPTER VI

RANK CORRELATION DATA

I. INTRODUCTION

In determining the correlation of the twelve WISC subtests with the Verbal Score, Performance Score, and Full Scale Score, Wechsler employed the product-moment method of correlation, and corrected his obtained r's for spuriousness by applying the formula recommended by McNemar.[68] In the present study, however, the non-normal nature of the parent population rendered invalid the application not only of product-moment correlation, but of McNemar's corrective formula as well. In coping with these difficulties, two alternate procedures were undertaken. First, correlation of subtest scores with Verbal, Performance, and Full Scale scores was accomplished by the employment of Kendall's method of rank correlation. In the second place, the computation of spuriously high tau's was prevented by correlating each subtest weighted score with the total score minus the subtest involved, rather than with the unaltered total score.

Each of the 57 superior subjects under consideration in this investigation was administered the ten regular

WISC subtests. As a result, the following 13 scaled or weighted scores were determined for each subject: (a) 5 Verbal subtest scores; (b) 5 Performance subtest scores; (c) a total Verbal score; (d) a total Performance score; (e) a Full Scale score. Total Verbal, total Performance, and Full Scale scores were then converted into Verbal, Performance, and Full Scale Intelligence Quotients respectively by reference to Table XI in the WISC manual. For the reader's information, the 13 weighted scores and the 3 Intelligence Quotients obtained for each of the 57 subjects are presented in Tables XXVI and XXVII, Appendix A.

The first step in the calculation of the rank correlation coefficients was the determination of the ranks of the 57 scaled scores for each of the 10 subtests. The method of ranking adopted was that of assigning to the highest of the 57 scores the rank of one, and to the lowest, the rank of 57. Where two or more individuals obtained the same weighted score, each was allotted the mean rank of the tied scores. The method of ranking the 57 Vocabulary weighted scores is indicated in Table XXVIII, Appendix B, while the ranks assigned to the 57 subjects on each of the subtests, and on total Verbal and total Performance, are furnished in Table XXIX, Appendix B.

The next procedure was that of subtracting in turn each of the 10 regular subtest scores from the total score obtained by a subject. This rendered 10 'partial totals' constituting the various combinations of 9 subtest scores. Each of the 10 groups of 57 scores obtained in this manner was then ranked following the aforementioned method. The resulting ranks allotted to the 57 subjects are to be found in Table XXX, Appendix B.

The total Verbal score of an individual was now diminished by each of his 5 Verbal subtest scores in turn. This gave scores which were designated as 'total Verbal minus Information', 'total Verbal minus Comprehension', 'total Verbal minus Arithmetic', 'total Verbal minus Similarities', and 'total Verbal minus Vocabulary'. The 57 subjects were then ranked according to the magnitude of these 'partial totals'. The 5 rankings are shown in Table XXXI, Appendix B.

Similar Performance 'partial totals' were determined, rendering scores labelled as 'total Performance minus Picture Completion', 'total Performance minus Picture Arrangement', 'total Performance minus Block Design', 'total Performance minus Object Assembly', and 'total Performance minus Coding'. The ranks assigned to the 57 subjects on the basis of these 5 combinations of Performance subtest scores are to be found in Table XXXII, Appendix B.

After the ranking of the subjects had been completed, the actual calculation of the rank correlation coefficients was carried out. This involved the calculation of a score 'S' whose magnitude was determined by the amount of agreement between a given pair of rankings of 57 subjects. The rank correlation coefficient 'tau_b' was then computed by applying the formula[69]

$$tau_b = \frac{S}{\sqrt{\frac{1}{2}n(n-1) - T}\ \sqrt{\frac{1}{2}n(n-1) - U}}$$

in which case n was 57, and T and U the corrections for ties in the two rankings. The computation of the rank correlation of the Vocabulary weighted scores with the sums of the 4 other Verbal weighted scores is carried out in detail in Appendix C. The procedure followed in testing the level of significance of the 'tau' just obtained is also illustrated in Appendix C.

II. RESULTS OF THE CORRELATION ANALYSIS

The obtained correlation of each of the ten subtest weighted scores with the total of the nine other subtest weighted scores is given in Table IX.

The coefficients of rank correlation of the 5 Verbal subtests with the total Verbal score, and of the 5 Performance subtests with the total Performance score are indicated, along with the levels of significance of these co-

efficients, in Tables X and XI.

TABLE IX. COEFFICIENTS OF RANK CORRELATION OF WISC SUBTEST SCORES WITH FULL SCALE SCORE MINUS THE SUBTEST

Subtest	Tau_b	Levels of Significance			
		.10	.05	.01	.001
Information	0.18	Yes	No	No	No
Comprehension	-0.04	No	No	No	No
Arithmetic	0.06	No	No	No	No
Similarities	0.08	No	No	No	No
Vocabulary	0.15	No	No	No	No
Picture Completion	0.16	No	No	No	No
Picture Arrangement	-0.04	No	No	No	No
Block Design	0.24	Yes	Yes	No	No
Object Assembly	-0.05	No	No	No	No
Coding	-0.20	Yes	Yes	No	No

TABLE X. COEFFICIENTS OF RANK CORRELATION OF WISC VERBAL SUBTEST SCORES WITH TOTAL VERBAL SCORE MINUS THE SUBTEST

Subtest	Tau_b	Levels of Significance			
		.10	.05	.01	.001
Information	0.52	Yes	Yes	Yes	Yes
Comprehension	0.23	Yes	Yes	No	No
Arithmetic	0.13	No	No	No	No
Similarities	0.32	Yes	Yes	Yes	Yes
Vocabulary	0.33	Yes	Yes	Yes	Yes

TABLE XI. COEFFICIENTS OF RANK CORRELATION OF WISC PERFORMANCE SUBTEST SCORES WITH TOTAL PERFORMANCE SCORE MINUS THE SUBTEST

Subtest	Tau_b	Levels of Significance			
		.10	.05	.01	.001
Picture Completion	0.17	Yes	No	No	No
Picture Arrangement	0.20	Yes	Yes	No	No
Block Design	0.25	Yes	Yes	Yes	No
Object Assembly	0.23	Yes	Yes	No	No
Coding	0.04	No	No	No	No

The totals of the 5 Verbal subtest scores were correlated with those of the 5 Perfomance subtest scores, resulting in a coefficient of rank correlation of -0.15. This coefficient was not significantly different from zero, however, at even the .10 level of confidence. For the purpose of comparing the magnitudes of two coefficients computed by altogether different methods, the product-moment correlation was now determined from the same data, and found to be -0.16. Once again the coefficient was not significantly different from zero, bearing out the fact that no relationship existed between total Verbal and total Performance scores for this sample of superior subjects.

The weighted scores obtained by 52 subjects on the Digit Span subtest were ranked, as were the 52 corresponding total Verbal and Full Scale scores. These three

rankings are presented in Table XXXIII, Appendix B. Since the Digit Span scores affected the calculation of neither the Verbal nor the Full Scale Intelligence Quotients of the 52 subjects, no correction for spuriousness was required in the computation of rank correlation coefficients involving Digit Span results. Consequently, the tau's were calculated directly between the Digit Span weighted scores and the totals of the 5 regular Verbal scores in one case, and the totals of the 10 subtest scores in the other. The coefficients of rank correlation of the Digit Span subtest with the total Verbal and with the Full Scale are to be found in Table XII.

TABLE XII. COEFFICIENTS OF RANK CORRELATION OF WISC DIGIT SPAN SUBTEST SCORES WITH TOTAL VERBAL AND FULL SCALE SCORES OBTAINED IN A SAMPLE OF 52 MENTALLY SUPERIOR CHILDREN

Correlation	Tau_b	Levels of Significance			
		.10	.05	.01	.001
Digit Span X Total Verbal	0.13	No	No	No	No
Digit Span X Full Scale	-0.04	No	No	No	No

III. DISCUSSION

Table IX indicates that the correlation between scores on each of the 10 subtests and the Full Scale was extremely low in all cases, ranging from 0.24 to -0.20. Only 2 of the

10 correlations were significantly different from zero at the .05 level, those of Block Design with Full Scale, and Coding with Full Scale. Of special note was the fact that there was a significant negative correlation between the Coding subtest scores and the Full Scale scores. This would suggest that in the present sample of Superior children there was a slight but significant tendency for the Coding scores to be low when the sum of the other subtest scores was high, and vice versa. The Block Design subtest score was evidently the best single predictor of the Full Scale score since this subtest correlated to a greater degree than any of the others with the sum of the 9 other subtest scores.

It may be seen in Table X that all of the 5 Verbal subtests with the single exception of Arithmetic correlated significantly with the sum of the 4 remaining Verbal subtests at the .05 level. The Information subtest, correlating .52 with the other Verbal subtests, was evidently the best predictor of the total Verbal score for this sample of Superior subjects.

On the whole, there was a lower degree of correlation between Performance subtest scores and total Performance than between Verbal subtest scores and total Verbal. Only 3 of the 5 Performance subtest scores, namely those of

Picture Arrangement, Block Design, and Object Assembly, correlated significantly with the sum of the 4 other Performance scores at the .05 level.

There was no significant relationship between total Verbal and total Performance scores in this sample, whereas in Wechsler's standardization group a correlation of about .60 was found to exist between these two sets of scores.[70] The lack of correlation of Verbal with Performance scores in the present sample may be partially explained by the fact that the range of Verbal and Performance Intelligence Quotients was only about 40 I.Q. points, while the WISC was standardized on a group that varied nearly 110 points in both Verbal and Performance Intelligence Quotients. Narrowing the I.Q. range in this manner and increasing the homogeneity of the group would logically tend to diminish the degree of relationship between any two sets of scores, whether they be total Verbal and total Performance, Verbal subtest and total Verbal, or Performance subtest and Full Scale.[71]

A comparison of the intercorrelation coefficients obtained in the present study with those determined by Wechsler at the 13½ year level revealed a number of similarities in the order of magnitude of the coefficients.[72] It was noted, for example, that in both the standardization

and the current samples: (a) of the 6 Verbal subtests, Information correlated highest with total Verbal; (b) the Digit Span and Arithmetic subtests showed the least correlation with total Verbal; (c) the correlations of the Block Design and Object Assembly subtests with total Performance were of greatest magnitude in the Performance area; (d) the positive relationship between Coding and total Performance was relatively weak; (e) the Information subtest correlated to a considerably greater degree with the Full Scale than did the Digit Span subtest; (f) of the 5 Performance subtests, Block Design showed the greatest correlation with the Full Scale.

Stacey and Levin found that total Verbal and total Performance correlated .21 and -.31 in samples of 44 morons and 26 borderline cases respectively[73], while Taylor obtained a rank correlation coefficient of .21 in a study involving 74 feeble-minded subjects.[74] The corresponding rank correlation coefficient obtained in the present investigation was not significantly different from zero, and appeared to compare more favorably with those derived in the Stacey and Levin and the Taylor studies than it did with Wechsler's product-moment coefficients of .56, .60, and .68 at three age levels of the standardization sample.[75]

IV. CONCLUSIONS

1. A considerably lower degree of correlation between WISC subtest scores and total Verbal, total Performance, and Full Scale scores was found in this sample of 57 Superior children than had previously been obtained for the WISC standardization sample of 2200 randomly chosen children.

2. Judging from the results obtained in this investigation and in those conducted by Stacey and Levin, and by Taylor, the degree of correspondence between total Verbal and total Performance scores obtained from intellectually homogeneous samples is very slight.

3. Block Design is possibly the strongest of the 11 subtests as a predictor of the Full Scale score of a superior child.

4. The scaled score obtained by a superior subject on Information likely affords the best means of predicting his total Verbal score.

5. Block Design appears to be a better measure than any other subtest of the abilities appraised on the Performance Scale.

CHAPTER VII

MEANS AND MEAN DIFFERENCES

I. THE MEANS

The subtest scaled scores obtained by the 57 superior subjects tested on the WISC are presented in Table XXVI, Appendix A. After grouping the scores on the 11 subtests into frequency distributions and choosing assumed means, the true means were computed by application of the formula[76]

$$M = A.M. + (\Sigma fx^1/N)i$$

The variability of these scores about their means was found by calculating the standard deviations of the 11 distributions. This was accomplished by the short method, the formula for which is[77]

$$S.D. = i\sqrt{\frac{\Sigma(fx^1)^2}{N} - c^2}$$

The range, mean, and S.D. of each of the 11 distributions of subtest scaled scores are shown in Table XIII.

The Verbal and Performance subtest scaled scores are totalled in columns (1) and (2) of Table XXVII, Appendix A, while Full Scale scores obtained by the 57 subjects appear in column (3). The ranges and means of these 3

distributions of scaled score totals are presented in Table XIV.

TABLE XIII. RANGES, MEANS, AND STANDARD DEVIATIONS OF WISC SUBTEST WEIGHTED SCORES OBTAINED BY 57 MENTALLY SUPERIOR CHILDREN

Subtest	Weighted Scores		
	Range	Mean	S.D.
Information	9 - 18	14.44	2.17
Comprehension	12 - 20	15.35	2.10
Arithmetic	10 - 19	14.93	2.25
Similarities	9 - 20	14.40	2.56
Vocabulary	9 - 18	13.81	1.84
Digit Span (N = 52)	8 - 19	13.35	2.53
Picture Completion	7 - 19	13.26	2.92
Picture Arrangement	6 - 17	12.70	2.28
Block Design	10 - 18	14.30	2.21
Object Assembly	9 - 19	13.04	2.29
Coding	8 - 20	12.77	2.65

TABLE XIV. RANGES AND MEANS OF WISC TOTAL VERBAL, TOTAL PERFORMANCE, AND FULL SCALE WEIGHTED SCORES OBTAINED BY 57 SUPERIOR CHILDREN

Group of Subtests	Weighted Scores	
	Range	Mean
Total Verbal (5)	57 - 90	72.93
Total Performance (5)	49 - 80	66.07
Full Scale (10)	127 - 165	139.00

The three WISC Intelligence Quotients obtained by each of the 57 subjects are to be found in columns (4), (5) and (6) of Table XXVII, Appendix A. Ranges, means, and standard deviations of the 3 distributions of Intelligence Quotients are shown in Table XV.

TABLE XV. RANGES, MEANS, AND STANDARD DEVIATIONS OF WISC VERBAL, PERFORMANCE, AND FULL SCALE INTELLIGENCE QUOTIENTS OBTAINED BY 57 SUPERIOR CHILDREN

Type of Quotient	WISC Intelligence Quotients		
	Range	Mean	S.D.
Verbal I.Q.	109 - 150	128.84	9.00
Performance I.Q.	99 - 142	122.51	9.35
Full Scale I.Q.	120 - 147	128.35	6.41

II. SIGNIFICANCE OF THE DIFFERENCES BETWEEN MEANS

In the measurement of intelligence, certain assumptions are necessary for the validation of the many mathematical procedures employed by those who construct the tests. According to Wechsler, one of the greatest contributions of Binet was his intuitive assumption that in the selection of the tests, it made little difference what tasks were used, provided that in some way they measured the child's general intelligence.[78]

This method of combining a variety of tests into a

single measure of intelligence presupposes a certain functional unity or equivalence between the tests.

> The functional equivalence of the test items, an assumption implicit not only in the Binet Scale but in any scale which is composed of a variety or pool of intellectual tasks, is absolutely necessary for the validation of the arithmetic employed in arriving at a final measure of intelligence. This arithmetic consists, first, of assigning some numerical value to every correct response; secondly, of adding the partial credits so obtained into a simple sum; and, thirdly, of treating equal sums as equivalent, irrespective of the nature of the test items which contribute to the total.[79]

The WISC scaled scores have been so derived as to provide, at each age and for each of the separate tests, a mean scaled score of 10 and a standard deviation of 3.[80] The Verbal and Performance Intelligence Quotients are dependent upon the sums of 5 Verbal and 5 Performance subtest scaled scores respectively. The Full Scale I.Q. is derived by adding the totals of the Verbal and the Performance scaled scores to obtain a Full Scale score, and then by referring to a table which converts these Full Scale scores into Full Scale Intelligence Quotients.[81] Thus each WISC subtest may be considered to contribute an equal amount to the determination of the Full Scale I.Q.

The means of the 10 subtest scaled scores obtained from a large, randomly-chosen sample should theoretically all have the value of 10, rendering total Verbal and total Performance mean scores of 50, a Full Scale mean score of

100, and mean Verbal, Performance and Full Scale Intelligence Quotients of 100. Differences occurring in the various subtest means, and in the means of the Verbal and Performance Intelligence Quotients should be small if they exist at all, and should be attributable only to errors in sampling.

In practice, it has been found that in samples of intellectually atypical subjects, such is rarely the case. These more homogeneous groups appear to do considerably better on certain of the subtests than on others, and mean scores may vary to the degree that one cannot attribute their differences solely to sampling errors. In order to test the hypothesis that certain of the WISC subtests offer less difficulty than others at the superior level, a statistical analysis of the differences between mean subtest scores was undertaken.

The significance of the difference between any pair of subtest scaled score means was determined by calculating the 't' or critical ratio associated with the difference, and by then referring to a t-table for the purpose of testing the magnitude of the calculated critical ratio. It was felt by members of the Department of Mathematics, University of Alberta, that the commonly employed t-formulas should not be applied with the current data, inasmuch as the use of these statistical methods presup-

posed certain conditions of normality, independence of means, and common variance that could not be assumed to exist. An alternate procedure, that of computing t's by comparison of scores in pairs, was suggested as the most appropriate. Accordingly, all critical ratios were calculated by applying the formula[82]

$$t = \frac{\bar{d}}{\sqrt{\frac{\sum(d-\bar{d})^2}{n(n-1)}}}$$

Table XVI(a) indicates the levels of significance of the differences between subtest mean scaled scores obtained in this study. Table XVI(b) is of a similar nature, but involves the Digit Span subtest, with N of 52 rather than of 57.

The difference between the mean Verbal and the mean Performance Intelligence Quotients obtained in this investigation was tested for significance by the employment of the Brookes and Dick formula. The actual difference of 6.33 I.Q. points was found to be significant at the .01 level.

III. DISCUSSION

The magnitude of the obtained subtest weighted or scaled scores varied from 6 up to the ceiling of 20, with the range in all cases being about 10 scaled score units.

TABLE XVI(a). DETERMINATION OF SIGNIFICANCE OF THE DIFFERENCES BETWEEN VARIOUS PAIRS OF WISC MEAN SCALED SCORES FOR THE SAMPLE OF 57 MENTALLY SUPERIOR SUBJECTS

Combination of Subtests	Mean Scaled Scores		Difference Between Means	Standard Error of Difference	Critical Ratio or 't'	Significance Level Attained
	First Subtest	Second Subtest				
Information-						
Comprehension	14.439	15.351	0.912	0.344	2.65	0.02
Arithmetic	14.439	14.930	0.491	0.356	1.38	
Similarities	14.439	14.404	0.035		<1	
Vocabulary	14.439	13.807	0.632	0.252	2.51	0.02
P. Completion	14.439	13.263	1.176	0.465	2.53	0.02
P. Arrangem't	14.439	12.702	1.737	0.438	3.97	0.01
Block Design	14.439	14.298	0.141		<1	
O. Assembly	14.439	13.035	1.404	0.457	3.07	0.01
Coding	14.439	12.772	1.667	0.515	3.24	0.01

TABLE XVI(a). (Continued)

Combination of Subtests	Mean Scaled Scores		Difference Between Means	Standard Error of Difference	Critical Ratio or 't'	Significance Level Attained
	First Subtest	Second Subtest				
Comprehension-						
Arithmetic	15.351	14.930	0.421		<2	
Similarities	15.351	14.404	0.947	0.381	2.49	0.02
Vocabulary	15.351	13.807	1.544	0.329	4.69	0.01
P. Completion	15.351	13.263	2.088		>3	0.01
P. Arrangem't	15.351	12.702	2.649		>3	0.01
Block Design	15.351	14.298	1.053	0.460	2.29	0.05
O. Assembly	15.351	13.035	2.316		>3	0.01
Coding	15.351	12.772	2.579		>3	0.01
Arithmetic-						
Similarities	14.930	14.404	0.526	0.441	1.19	
Vocabulary	14.930	13.807	1.123	0.353	3.18	0.01
P. Completion	14.930	13.263	1.667	0.493	3.38	0.01
P. Arrangem't	14.930	12.702	2.228		>3	0.01
Block Design	14.930	14.298	0.632	0.356	1.77	
O. Assembly	14.930	13.035	1.895	0.450	4.21	0.01
Coding	14.930	12.772	2.158		>3	0.01

TABLE XVI(a). (Continued)

Combination of Subtests	Mean Scaled Scores		Difference Between Means	Standard Error of Difference	Critical Ratio or 't'	Significance Level Attained
	First Subtest	Second Subtest				
Similarities-						
Vocabulary	14.404	13.807	0.597	0.336	1.78	
P. Completion	14.404	13.263	1.141	0.479	2.38	0.05
P. Arrangem't	14.404	12.702	1.702	0.495	3.44	0.01
Block Design	14.404	14.298	0.106		< 1	
O. Assembly	14.404	13.035	1.369	0.507	2.70	0.01
Coding	14.404	12.772	1.632	0.545	2.99	0.01
Vocabulary-						
P. Completion	13.807	13.263	0.544		< 2	
P. Arrangem't	13.807	12.702	1.105	0.406	2.72	0.01
Block Design	13.807	14.298	0.491		< 2	
O. Assembly	13.807	13.035	0.772	0.428	1.80	
Coding	13.807	12.772	1.035	0.484	2.14	0.05

TABLE XVI(a). (Continued)

Combination of Subtests	Mean Scaled Scores		Difference Between Means	Standard Error of Difference	Critical Ratio or 't'	Significance Level Attained
	First Subtest	Second Subtest				
P. Completion-						
P. Arrangem't	13.263	12.702	0.561	0.459	1.22	
Block Design	13.263	14.298	1.035	0.417	2.48	0.02
O. Assembly	13.263	13.035	0.228		<1	
Coding	13.263	12.772	0.491		<2	
P. Arrangem't-						
Block Design	12.702	14.298	1.596	0.410	3.89	0.01
O. Assembly	12.702	13.035	0.333		<2	
Coding	12.702	12.772	0.070		<1	
Block Design-						
O. Assembly	14.298	13.035	1.263	0.347	3.64	0.01
Coding	14.298	12.772	1.526	0.467	3.27	0.01
O. Assembly-						
Coding	13.035	12.772	0.263		<1	

TABLE XVI(b). DETERMINATION OF SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE DIGIT SPAN MEAN SCALED SCORE AND THE MEAN SCALED SCORES OF THE REMAINING WISC SUBTESTS FOR 52 MENTALLY SUPERIOR SUBJECTS

Subtest Paired With Digit Span	Mean Scaled Scores		Difference Between Means	Standard Error of Difference	Critical Ratio or 't'	Significance Level Attained
	Digit Span	Other Subtest				
Information	13.346	14.308	0.962	0.410	2.35	0.05
Comprehension	13.346	15.365	2.019		>3	0.01
Arithmetic	13.346	14.865	1.519	0.436	3.48	0.01
Similarities	13.346	14.250	0.904	0.479	1.89	
Vocabulary	13.346	13.750	0.404		<2	
P. Completion	13.346	13.154	0.192		<1	
P. Arrangem't	13.346	12.635	0.711	0.518	1.37	
Block Design	13.346	14.192	0.846	0.504	1.68	
O. Assembly	13.346	13.019	0.327		<2	
Coding	13.346	12.885	0.461		<2	

Subtest means ranged from 12.70 to 15.35, indicating that the group's achievement was above average on all of the eleven WISC subtests. Highest mean scores were obtained in Comprehension and Arithmetic, while the Picture Arrangement and Coding tests apparently offered the greatest difficulty, since their mean scores were the lowest of the eleven.

The standard deviations of the subtest weighted scores were commonly between 2.1 and 2.3 score units, with extreme values of 1.8 and 2.9 occurring in the Vocabulary and Picture Completion distributions respectively. Standard deviations of the subtest score distributions were on the whole rather less than those obtained in the standardization research, where sigma values of 2.6 to 3.2 were common.[83] The fact that the variability of scores in the present sample is not as great as it was in Wechsler's is not surprising, since a more homogeneous group is being examined.

The range of both Verbal and Performance Intelligence Quotients obtained in the study was just over 40 points of I.Q., with the Verbal upper and lower limits being about 9 points higher than the corresponding Performance limits. The mean Verbal I.Q. of the 57 subjects closely approximated their mean Full Scale I.Q., while both of these means exceeded the average of the Performance Intelligence Quotients by nearly $6\frac{1}{2}$ points. This finding is in agreement

with one reported by Wechsler.[84]

It should be noted that although individual Verbal and Performance Intelligence Quotients were as low as 109 and 99 respectively, the least Full Scale I.Q. was nevertheless 120. This peculiarity is explained by the fact that positive and negative discrepancy scores between Verbal and Performance Intelligence Quotients were quite large in several instances. To cite an example, the subject who scored 109 I.Q. on the Verbal Scale obtained a Performance quotient of 140, resulting in a Full Scale I.Q. of 126. It will be recalled that all subjects considered in this investigation attained or exceeded a Full Scale I.Q. of 120, regardless of their achievement on the Verbal and Performance Scales separately.

The standard deviations of both the Verbal and Performance I.Q. distributions were considerably less than those obtained in Wechsler's research, being only about 9 as compared with 15 I.Q. points.[85] This diminished variability of I.Q. is understandable, inasmuch as only superior subjects were considered in the current study.

Many of the observed means were found to be significantly different at the .01 level of confidence. The writer was interested in knowing the probability that given differences or greater positive ones might have arisen from sampling errors when the true differences were zero. Where

two means were significantly different at the .02 level, the probability that the smaller mean was equal to or exceeded the larger in the universe was one-half of .02, or .01, since a one-tail rather than a two-tail test was involved.[86]

Accordingly, a number of conclusions were drawn regarding the differing magnitudes of the subtest means obtained in this study. These findings are listed below.

1. The mean scores obtained by the superior subjects on Comprehension and Arithmetic were significantly greater at the .01 level than their mean scores on Vocabulary, Digit Span, Picture Completion, Picture Arrangement, Object Assembly, and Coding. The Comprehension mean, in addition, was found to be greater than the Information and Similarities means at this level of confidence.

2. The mean scores obtained on Information and Similarities were significantly greater at the .01 level than the mean scores on Picture Arrangement, Object Assembly, and Coding. The Information mean also exceeded those of Vocabulary and Picture Completion at this level.

3. The mean score obtained on Block Design was significantly greater at the .01 level than the mean scores computed for all of the remaining Performance subtests.

4. No Verbal subtest mean score was significantly greater at the .01 level than that obtained on Block

Design.

5. Picture Arrangement, Object Assembly, and Coding means were each exceeded by the means of at least five other subtests at the .01 level of confidence.

6. The Digit Span mean score was not significantly greater at the .01 level than the mean scores obtained on any of the five Performance subtests.

IV. SUMMARY AND CONCLUSIONS

On the assumption that the WISC subtests are functionally equivalent, and that all contribute by equal amounts to a measure of the global capacity of a child, a statistical analysis was undertaken in an effort to determine whether or not certain of the subtests offer less difficulty than others for superior subjects. The conclusions to be drawn are as follows:

1. For this sample of 57 superior children, the Comprehension and Arithmetic subtests were less difficult than the majority of the remaining WISC subtests;

2. The Information, Similarities, and Block Design subtests offered less challenge than three of the Performance subtests: Picture Arrangement, Object Assembly, and Coding;

3. The Block Design subtest was found by the group to be less complex than any of the other Performance sub-

tests;

4. Picture Arrangement, Object Assembly, and Coding were all more difficult than at least five of the remaining ten WISC subtests;

5. The Performance Scale was, on the whole, significantly more difficult for this group of superior subjects than was the Verbal Scale.

CHAPTER VIII

THE DISCREPANCY SCORE SURVEY

Seashore has defined a WISC "discrepancy score" as the numerical difference between a subject's Verbal I.Q. and his Performance I.Q.[87] In the calculation of these scores, Performance I.Q. is always subtracted from Verbal I.Q., making possible both positive and negative discrepancies.[88] Thus if Subject A has a Verbal I.Q. of 135 and a Performance I.Q. of 112, his discrepancy score is +23. On the other hand, Subject B with a Verbal I.Q. of 119 and a Performance I.Q. of 127 has a discrepancy score of -8.

The writer wished to compare the discrepancy scores obtained in the present investigation with those found both in Taylor's sample of 74 feeble-minded subjects,[89] and in the standardization study. The object of this survey was to determine whether or not a definite trend exists in the nature of the distributions of discrepancy scores obtained in samples of atypical subjects.

The discrepancy scores obtained in the standardization study have been estimated by considering alternately Figure 1 and Table III included in Seashore's publication.[90] Although the writer could not determine the frequencies at

the various levels of discrepancy with perfect accuracy due to the lack of precise data presented in the article, nevertheless it was felt that the frequencies arrived at in column (2) of Table XIX are very nearly authentic. For the convenience of the reader, both Figure 1 and Table III accompanying Seashore's article have been reproduced below in Figure 2 and Table XVII.

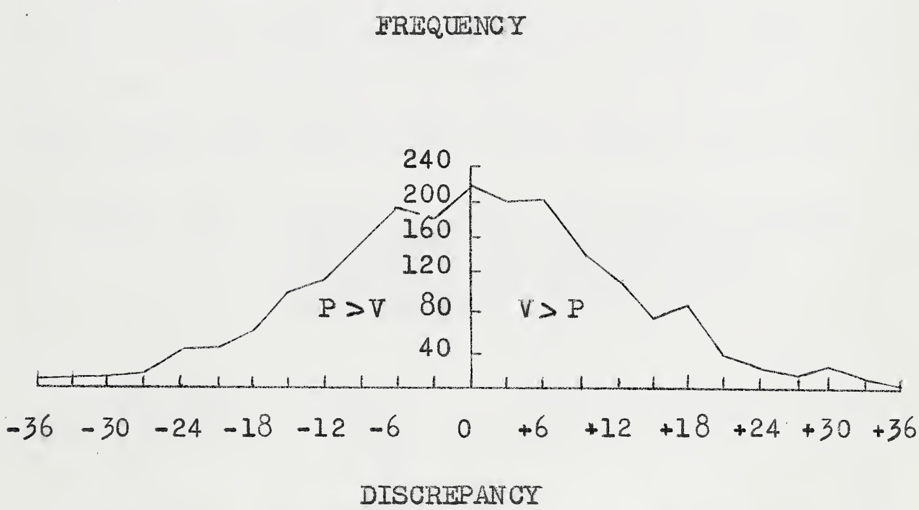

Figure 2. Distribution of Differences between Verbal and Performance IQs for 2200 cases, age 5-15, from standardization sample on the Wechsler Intelligence Scale for Children.

(This is a reproduction of Fig. 1 accompanying the article "Differences between Verbal and Performance IQs on the WISC".)[91]

An examination of Figure 2 indicates that the discrepancy scores for the 2200 subjects in the standardiza-

tion sample were grouped in intervals of 3 before the construction of the frequency polygon was undertaken. Midpoints of these intervals are seen to be 0, ±3, ±6, ±9, ... ±36. The frequency in the central interval -1.5 to +1.5 may be estimated from the polygon to be slightly more than 220. Table XVII discloses that the number of subjects obtaining absolute discrepancy scores of 0 and 1 was actually 86 plus 140, i.e. 226. This frequency of 226 has been placed in its appropriate position in Table XVIII, and the remainder of this table was drawn up by working alternately from Figure 2 to Table XVII in the manner indicated. The

TABLE XVII. DISTRIBUTION OF DISCREPANCY SCORES (REGARDLESS OF SIGN) FOR THE STANDARDIZATION SAMPLE OF 2200 CHILDREN TESTED ON THE WISC

Size of V and P Differences	Number of Cases	Size of V and P Differences	Number of Cases
35 +	14	8	112
30 - 34	36	7	163
25 - 29	68	6	114
20 - 24	140	5	131
15 - 19	284	4	132
14	68	3	124
13	74	2	144
12	83	1	140
11	88	0	86
10	82		
9	117	Total N	2200

(Table XVII is a reproduction of the first two columns of Table III accompanying H.G. Seashore's article "Differences between Verbal and Performance IQs on the Wechsler Intelligence Scale for Children.")92

total frequency obtained by this procedure was 2195, considered by the writer to be reasonably close to the value of 2200.

Examinations of Figure 2 and Table XVII led to estimated positive and negative discrepancy score frequencies for the 2200 subjects. These estimated frequencies are presented in Table XVIII.

TABLE XVIII. DISTRIBUTION OF POSITIVE AND NEGATIVE DISCREPANCY SCORES OF 2200 SUBJECTS TESTED ON THE WISC, AS DERIVED BY INSPECTION OF FIGURE 2 AND TABLE XVII

Class-interval	Estimated Frequency	Class-interval	Estimated Frequency
-37.5 to -34.5	5	1.5 to 4.5	210
-34.5 to -31.5	8	4.5 to 7.5	210
-31.5 to -28.5	12	7.5 to 10.5	152
-28.5 to -25.5	20	10.5 to 13.5	122
-25.5 to -22.5	42	13.5 to 16.5	80
-22.5 to -19.5	46	16.5 to 19.5	92
-19.5 to -16.5	64	19.5 to 22.5	42
-16.5 to -13.5	109	22.5 to 25.5	29
-13.5 to -10.5	123	25.5 to 28.5	16
-10.5 to -7.5	159	28.5 to 31.5	22
-7.5 to -4.5	198	31.5 to 34.5	12
-4.5 to -1.5	190	34.5 to 37.5	6
-1.5 to 1.5	226		
		Total N	2195

The frequencies on class-intervals of three presented in Table XV have been regrouped on class-intervals of nine in column (2) of Table XIX. The discrepancy scores obtained by Taylor for 74 feeble-minded subjects have been

placed in column (1) of the same table, while the corresponding scores found in the present investigation appear in column (3). The frequencies of the three sets of discrepancy scores have been grouped on comparatively wide intervals of nine to facilitate the construction of unsmoothed percentage frequency polygons.

TABLE XIX. DISTRIBUTIONS OF DISCREPANCY SCORES FOR SAMPLES OF 74 FEEBLE-MINDED, 2200 RANDOMLY-SELECTED, AND 57 MENTALLY SUPERIOR CHILDREN TESTED ON THE WISC

Class-interval	Midpoint of Interval	Obtained Frequencies		
		(1) Feeble-minded Group	(2) Randomly-selected Group	(3) Superior Group
-40.5 to -31.5	-36	0	13	0
-31.5 to -22.5	-27	2	74	1
-22.5 to -13.5	-18	10	219	2
-13.5 to -4.5	-9	21	480	8
-4.5 to 4.5	0	23	626	16
4.5 to 13.5	9	13	484	14
13.5 to 22.5	18	4	214	9
22.5 to 31.5	27	1	67	3
31.5 to 40.5	36	0	18	4
Total Frequencies		74	(2195)	57

The next procedure was to convert the frequencies on the various intervals into percentages of the total populations being considered. The percentage frequencies obtained appear in Table XX.

TABLE XX. PERCENTAGE FREQUENCY DISTRIBUTIONS OF DISCREPANCY SCORES FOR SAMPLES OF 74 FEEBLE-MINDED, 2200 RANDOMLY-SELECTED, AND 57 MENTALLY SUPERIOR CHILDREN TESTED ON THE WISC

Class-interval	Midpoint of Interval	Obtained Percentage Frequencies		
		(1) Feeble-minded Group	(2) Randomly-selected Group	(3) Superior Group
-40.5 to -31.5	-36	0.0%	0.6%	0.0%
-31.5 to -22.5	-27	2.7%	3.4%	1.8%
-22.5 to -13.5	-18	13.5%	10.0%	3.5%
-13.5 to -4.5	-9	28.4%	21.8%	14.0%
-4.5 to 4.5	0	31.1%	28.5%	28.1%
4.5 to 13.5	9	17.6%	22.0%	24.6%
13.5 to 22.5	18	5.4%	9.7%	15.8%
22.5 to 31.5	27	1.4%	3.0%	5.3%
31.5 to 40.5	36	0.0%	0.8%	7.0%
Totals		100.1%	99.8%	100.1%

Due to the procedure of rounding off the percentage frequencies to one place of decimals, the sums of columns (1) and (3) in Table XX are slightly more than 100%. The N obtained by estimating the frequencies for the standardization sample was 2195, a value somewhat below the actual N of 2200. This accounts for the fact that the sum of the percentage frequencies in column (2) is only 99.8%.

The percentage frequencies appearing in columns (1), (2), and (3) of Table XX were then plotted on the same axes. Three comparable frequency polygons were constructed,

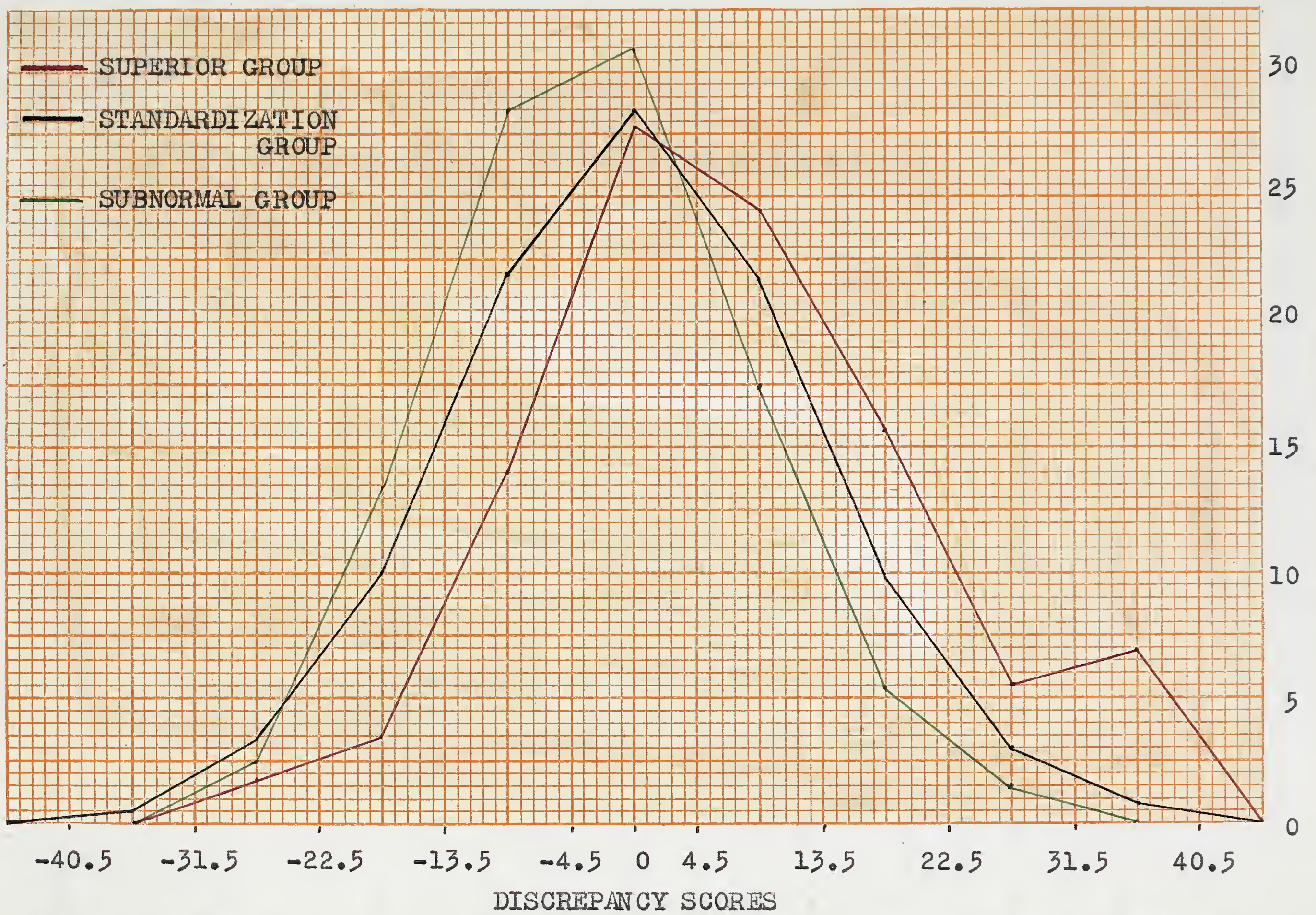

Fig. 3: Distributions of Discrepancy Score Percentage Frequencies for Three Samples of Children Tested on the WISC

and these constitute Figure 3.

DISCUSSION

Table XIX reveals that the frequency of the discrepancy scores is greatest in the central interval of each of the three distributions under consideration. These central intervals have midpoints of zero, suggesting that discrepancy scores tend to assume a value near zero regardless of whether the sample has been randomly chosen, or has been composed only of intellectually atypical subjects. It should be noted, however, that the frequencies in column (2) of Table XIX are far more symmetrically distributed about the central interval than are the frequencies in either of columns (1) or (3). Positive discrepancy scores for the subnormal group tend to diminish in frequency more rapidly as magnitude increases than do the corresponding negative scores. This trend is quite the opposite in the superior group, for the number of positive scores greatly outweighs the number of negative scores. In the third group, viz. the standardization sample, the rate of decrease in frequency with increasing absolute magnitude is practically identical for both positive and negative discrepancies.

Table XX bears out these findings. The percentage of

the frequencies within the central interval is practically constant for the three samples, varying only from values of 28% to 31%. However, whereas 44% of the discrepancy scores are less than -4.5 for the mentally deficient group, the proportion decreases to 36% for the randomly-selected group, and to only 19% for the sample of superior subjects. On the other hand, the percentage of the discrepancy scores which are greater than 4.5 increases from 24% for the subnormal group to 36% for the standardization sample and to 52% for the superior group. Thus there is apparently a tendency for the percentage of positive discrepancy scores to vary directly with the mean WISC Full Scale I.Q. of the sample.

The unsmoothed frequency polygons in Figure 3 indicate this tendency pictorially. The polygon representing the distribution of the subnormal group's discrepancy scores is skewed positively or to the right, while that of the superior group is skewed negatively. The distribution of discrepancy scores for the standardization sample differs from the two other distributions in that it is very nearly symmetric in form.

The aforementioned findings are in harmony with those reported by Wechsler for the Bellevue Scale standardization sample.[93] He found that of the 41 subjects in the 10 - 16

age group who attained I.Q. scores of more than 120 on the Wechsler-Bellevue, 82.9% had greater Verbal Intelligence Quotients than Performance Intelligence Quotients. The proportion of cases obtaining positive discrepancy scores decreased to 43.9% in the average group, and to 25.0% in the subnormal group, strongly indicating the same trend found in the present study.

CONCLUSIONS

1. The nature of the distribution of discrepancy scores obtained by a group of children tested on the WISC appears to depend upon the level of mental ability of the group. A highly intelligent group apparently tends to obtain a greater proportion of positive than of negative discrepancy scores, while the reverse tendency seems to be indicated for a mentally inferior group.

2. Although the standardization procedures for the WISC were designed to result in a difference of zero points of I.Q. between the averages of Verbal and Performance intelligence quotients, the mean I.Q. of a group appears to determine whether or not this desired difference of zero points will be obtained. It seems likely that the more the mean Full Scale I.Q. of an intellectually homogeneous group deviates from 100, the greater will be the difference between the means of the Verbal and the Performance Intelligence Quotients of the group.

CHAPTER IX

ADEQUACY OF SUBTEST CEILINGS

One of the chief criticisms of the Wechsler-Bellevue Scale is that it does not measure effectively the degree of superiority of gifted subjects.[94] Wechsler is well aware of this criticism and has responded by stating that it is not possible to measure general intelligence beyond a certain level.

> IQ's above 130, whether on the Bellevue or any other tests, do not represent superiority in intelligence level, but special aptitude in intellectual ability.[95]

At the same time, however, he has attempted to pacify the critics by extending the range of the Verbal part of the Wechsler-Bellevue Scale at the upper levels, and by extending the maximum score possible on them from 17 to 18. This move was designed to furnish more discriminating scores for Intelligence Quotients above 120.[96]

Little has been published to date on the adequacy of the WISC as a measure of the degree of intellectual superiority among children. The question apparently remains unanswered as to whether or not this more recent test of Wechsler's has eliminated the aforementioned basis on which the Bellevue Scale was criticized. In this light, consequently, the writer analyzed the scores obtained by

the 66 subjects of potentially superior intelligence who were tested in the present investigation.

Fifty-seven of these 66 testees earned the superior grading defined by Wechsler, 120 I.Q. on the Full Scale. Of the 57, no less than 11 obtained perfect raw scores on at least one subtest, while 2 achieved maximum allowable scores on 2 of the subtests. Of the 9 subjects who failed to score 120 I.Q. on the Full Scale, one obtained perfect raw scores on 2 subtests, and at the same time earned the maximum scaled or weighted rating of 20 points on a third subtest.

Perfect raw scores were obtained on 4 of the 11 WISC subtests administered in the investigation. Table XXI offers information as to the number, sex, and range in chronological age of the testees who supplied correct responses to all of the items, and who in addition earned the maximum time bonuses allowed, on the 4 subtests involved.

TABLE XXI. NUMBER, SEX, AND RANGE IN CHRONOLOGICAL AGE OF SUBJECTS OBTAINING PERFECT RAW SCORES ON VARIOUS WISC SUBTESTS

WISC Subtest	Number of Subjects			Range in Chronological Age
	Male	Female	Total	
Arithmetic	8	2	10	12-2 to 14-7
Digit Span	2	1	3	14-1 to 14-10
Object Assembly	1	0	1	12-6
Block Design	1	0	1	14-4

The male frequencies in Table XXI include the testee mentioned above who did not achieve a Full Scale I.Q. rating of 120, but who nevertheless obtained perfect raw scores on both the Arithmetic and Digit Span subtests. This subject had a chronological age of 14-2, and Verbal, Performance, and Full Scale Intelligence Quotients of 126, 99, and 115 respectively.

The various I.Q. ratings of the 11 Superior subjects who achieved perfect scores on one or more of the subtests are presented in Table XXII.

TABLE XXII. SEX AND WISC I.Q. RATINGS OF SUPERIOR SUBJECTS OBTAINING PERFECT RAW SCORES ON ONE OR MORE WISC SUBTESTS

Subject	Sex	Subtests Involved	WISC I.Q. Ratings		
			Verbal	Performance	Full Scale
A	F	Arithmetic	150	135	147
B	M	Arithmetic Block Design	142	142	146
C	M	Arithmetic	138	132	138
D	M	Arithmetic	135	122	132
E	M	Arithmetic	131	124	131
F	M	Object Assembly	126	125	128
G	M	Arithmetic	134	117	128
H	M	Arithmetic Digit Span	142	106	127
I	F	Arithmetic	123	127	127
J	M	Arithmetic	126	114	123
K	F	Digit Span	125	113	121

The means of the 3 WISC Intelligence Quotients of the 11 subjects involved are shown in Table XXIII, along with

comparable means for the 46 superior subjects who did not obtain perfect scores on any of the subtests.

TABLE XXIII. MEAN I.Q. RATINGS OF SUPERIOR SUBJECTS WHO DID AND DID NOT OBTAIN PERFECT RAW SCORES ON ONE OR MORE WISC SUBTESTS

Class of Subjects	N	Mean WISC I.Q. Ratings		
		Verbal	Performance	Full Scale
With perfect raw scores	11	133.8	123.4	131.6
Without perfect raw scores	46	127.7	122.3	127.6
Total N	57	128.8	122.5	128.3

Table XXIV has been designed to indicate the maximum raw scores obtained by the 57 subjects at different age levels. Although the data include the scores of only one testee in each of age groups 7, 8, 10, and 15, the frequencies are greater at the remaining age levels, and a trend toward higher scores with increasing age is noticeable.

The lone seven-year-old was not administered the Digit Span subtest, and the Coding A test result which he obtained was not comparable to the Coding B raw scores obtained by the testees of eight years and over.

Further analysis of the subtest results indicated that 6 subjects obtained raw scores sufficiently high to entitle them to subtest scaled scores of 20. A description of the

TABLE XXIV. HIGHEST RAW SCORES OBTAINED ON WISC SUBTESTS BY 57 SUPERIOR CHILDREN ACCORDING TO AGE

WISC Subtest	Perfect Raw Score	Highest Raw Score Obtained at Age 7	8	9	10	11	12	13	14	15
Information	30	16	12	20	16	24	25	27	27	24
Comprehension	28	15	12	19	15	24	24	24	25	24
Arithmetic	16	9	7	13	13	15	16	16	16	14
Similarities	28	20	11	14	12	19	23	24	23	18
Vocabulary	80	37	35	46	44	51	55	65	61	57
Digit Span	17	..	10	15	15	14	15	16	17	13
Picture Completion	20	11	15	16	11	17	18	19	17	17
Picture Arrangement	57	20	27	42	38	44	48	46	48	45
Block Design	55	34	39	40	40	50	49	53	55	36
Object Assembly	34	16	27	25	31	33	34	32	30	30
Coding	93	..	22	44	45	65	67	70	88	56
Number of Testees:		1	1	6	1	11	13	16	7	1

subjects who obtained these maximum scaled scores is offered in Table XXV.

TABLE XXV. SEX, AGE, AND RAW SCORES OF SUBJECTS EARNING SCALED SCORE RATINGS OF 20 POINTS ON VARIOUS WISC SUBTESTS

Subject	Sex	Age	Subtest Involved	Subtest Raw Score: Obtained	Required for 20 points
(1)	M	14-11	Coding	88	84-93
(2)	M	7-10	Similarities	20	15-28
(3)	F	13-11	Similarities	24	24-28
(4)	M	12- 2	Similarities	23	23-28
(5)	M	11- 4	Comprehension	24	24-28
(6)	M	14- 2	Coding	81	81-93

It may be pointed out here that Subject (6) in Table XXV failed to fulfill the requirements for the Superior grading as his Full Scale I.Q. was found to be only 115. He is the testee, incidentally, to whom reference has already been made.

DISCUSSION

Table XXI furnishes evidence to the effect that several of the WISC subtests are not supplied with material difficult enough to tax the mental potentialities of a substantial proportion of either Superior or Bright Normal subjects. One can hardly dispute the fact that a testee who has earned a perfect score on any of the 11 subtests may well have been capable of supplying correct responses to more difficult test items.

The Arithmetic subtest appears to be weakest in this respect since 9 of 57 subjects, or approximately 15% of the sample, were able to supply correct solutions to all of the 16 problems. The youngest testee who was capable of earning a perfect score in Arithmetic was just over 12 years of age, indicating that even Problem 16 is not too challenging for children who experience little difficulty in mathematics.

The writer feels that the Arithmetic and Digit Span subtests in particular will not necessarily assist in dif-

ferentiating between the mental abilities of two subjects who possess superior intelligence. To a lesser degree, the Object Assembly and Block Design subtests are subject to the same criticism.

An examination of Table XXI leads to the hypothesis that only six of the WISC subtests, viz. Information, Comprehension, Similarities, Vocabulary, Picture Arrangement, and Coding, have ceilings of sufficient height to allow for the complete testing of the Bright Normal or Superior subject. While not one of the current testees was able to achieve a perfect raw score of 20 on Picture Completion, the score of 19 obtained by a 13 year-old occasions some wonder in the mind of the writer that this subtest has not been completely mastered by older subjects.

The entries in Table XXIII imply that subjects who have realized perfect raw scores on any subtest possess somewhat higher Verbal and Full Scale Intelligence Quotients than subjects who have not been able to secure a maximum allowable score. The frequencies, however, are not large, and the finding is probably inconclusive.

Wechsler's method of equating all raw scores exceeding a certain value to the maximum scaled score of 20 appears to be open to criticism. This scheme does not allow for differentiation between the abilities of two subjects who have done especially well on certain of the subtests. To

cite an example, Subject Two described in Table XXV obtained the amazingly high raw score of 20 on the similarities subtest at the age of only 7 years, 10 months. The conversion table on page 35 of the WISC manual discloses that this subject has earned the maximum scaled score of 20. This is also the scaled score which he would have obtained had his raw score been as low as 15 or as high as 28.

Although Wechsler in his WISC conversion tables has provided for those subjects who earn average or below average raw scores, it appears that he has not concentrated too much attention on the potential accomplishments of superior children. In the present instance, the writer feels that Subject Two has earned a scaled score of possibly 22 to 24 points rather than the score of 20 which must be assigned to him. In a similar manner, Subject One in Table XXII seems to be somewhat hindered by the fact that he can earn no more than 20 points for his efforts on the Coding subtest.

Wechsler was aware that this type of weakness existed in the parent Wechsler-Bellevue Scale, and remedied the situation by raising the maximum scaled scores from 17 to 18.[97] He may feel the need in a future revision of the WISC to make some alteration of a similar nature.

CONCLUSIONS

1. There is concrete evidence to indicate that several of the WISC subtests do not possess sufficiently high ceilings to measure fully the mental abilities of certain gifted adolescents. The Arithmetic and Digit Span subtests appear to be weakest in this respect, since in the current study perfect raw scores were obtained most frequently on these tests. It was found that subjects were also capable of completely mastering the Object Assembly and Block Design subtests.

2. Only six of the WISC subtests, namely Information, Comprehension, Similarities, Vocabulary, Picture Arrangement, and Coding appear to have adequate ceilings. The Maze test was not administered in this study.

3. On certain of the subtests, the maximum scaled score of 20 appears to hinder the measurement of differences between the mental abilities of a few superior subjects. In the present investigation it was felt that the respective ranges of the raw scores assigned to scaled scores of 20 for the Coding and Similarities subtests were too broad, and that scaled scores above 20 should rightfully be introduced in these instances.

CHAPTER X

CONCLUSIONS AND RECOMMENDATIONS

I. CONCLUSIONS

This study was undertaken in an effort to gather information concerning the achievement of mentally superior children on a relatively new intelligence scale, the WISC. The scores of 57 subjects who attained or exceeded a Full Scale I.Q. of 120 were examined individually and collectively for the purpose of ascertaining their exact nature.

A summary of the conclusions drawn from this investigation follows:

1. It is evidently not possible to predict with accuracy the WISC Full Scale I.Q. of a mentally superior subject by reference to his achievement on any one of the subtests;

2. Decreasing the intellectual heterogeneity of a group of subjects to be tested on the WISC will likely result in a material lowering of all part-whole correlations associated with this scale;

3. Superior subjects have a tendency to obtain higher Verbal Intelligence Quotients than Performance Intelligence Quotients on the WISC;

4. Highly intelligent children tend to obtain above-average scaled scores on every WISC subtest;

5. Subjects of superior intelligence tend to experience less difficulty on the Comprehension and Arithmetic subtests than on the majority of the remaining subtests;

6. There is an apparent tendency for the percentage of positive discrepancy scores to vary directly as the mean WISC Full Scale I.Q. of an intellectually homogeneous group;

7. It appears that of the 11 WISC subtests administered in this investigation, only Information, Comprehension, Similarities, Vocabulary, Picture Arrangement, and Coding have ceilings of sufficient height to permit the complete testing of certain superior subjects;

8. The procedure adopted by Wechsler of equating numerous raw scores to an uppermost scaled score of 20 appears to be open to criticism inasmuch as it occasionally hinders the precise measurement of individual differences.

II. RECOMMENDATIONS AND SUGGESTIONS FOR FUTURE STUDIES

1. The number of studies conducted with mentally superior subjects is not great. Much has yet to be accomplished in this field.

2. The writer feels that the present investigation is one of the first of its kind. Some of the results obtained for this sample of 57 subjects are novel and unexpected. Future studies may either bear out these findings or disprove them.

3. A distribution of superior Intelligence Quotients is undeniably non-normal, a fact which renders invalid the use of Pearson's 'r'. The Kendall method of rank correlation employed in this study was found to be somewhat cumbersome for N as large as 57, chiefly because of the influence of the numerous tied ranks on the complexity of the mathematical procedures involved.

4. It is felt that a statistical analysis should be made in the near future of the results obtained by the 179 superior subjects who constituted the upper section of the standardization sample for the WISC.

REFERENCES

1. Peterson, Joseph. *Early Conceptions and Tests of Intelligence*, World Book Company, Yonkers-on-Hudson, New York, and Chicago, 1925, p. 257.

2. *Ibid*., p. 258.

3. *Ibid*., p. 267.

4. Chronbach, Lee J. *Essentials of Psychological Testing*, Harper and Brothers, New York, 1949, p. 106.

5. Peterson, Joseph. *op. cit*., p. 196.

6. *Ibid*., p. 237.

7. Richards, T. W. *Modern Clinical Psychology*, McGraw-Hill Book Company, New York and London, 1946, p. 53.

8. Wechsler, David. *Manual for the Wechsler Intelligence Scale for Children*, The Psychological Corporation, New York, 1949, p. 1.

9. Wechsler, David. *The Measurement of Adult Intelligence*, The Williams and Wilkins Company, Baltimore, 3rd Ed., 1944, p. 11.

10. *Ibid*.

11. American Association for Gifted Children. *The Gifted Child*, D. C. Heath and Company, Boston, 1951, p.v.

12. *Ibid*., p. 185.

13. *Ibid*., p. 4.

14. Louttit, C. M. *Clinical Psychology*, Harper and Brothers, New York and London, 1947, p. 319.

15. American Association for Gifted Children. *op. cit*., p. 204.

16. Peterson, Joseph. *op. cit*., p. 287

17. Chronbach, Lee J. *op. cit.*, p. 141

18. *Ibid.*, p. 158.

19. Wechsler, David. *The Measurement of Adult Intelligence*, *op. cit.*, p. 147.

20. Seashore, Harold and others. "The Standardization of the Wechsler Intelligence Scale for Children", *Journal of Consulting Psychology*, XIV, 1950, pp. 107-109.

21. Seashore, Harold G. "Differences between Verbal and Performance IQ's on the Wechsler Intelligence Scale for Children", *Journal of Consulting Psychology*, XV, 1951, pp. 62-67.

22. Wechsler, David. *op. cit.*, p. 125.

23. *Ibid.*, p. 126.

24. *Ibid.*

25. Stacey, C. L. and Levin, Janice. "Correlation Analysis of Scores of Subnormal Subjects on the Stanford-Binet and the WISC", *American Journal of Mental Deficiency*, LV, 1951, pp. 590-597.

26. Taylor, L. B. "An Investigation into the Nature of Results Obtained on the WISC by Mentally Deficient Children". (Unfinished master's thesis, University of Alberta).

27. Wechsler, David. *op. cit.*, p. 37.

28. *Ibid.*, p. 38.

29. *Ibid.*, p. 39.

30. *Ibid.*, p. 34.

31. *Ibid.*, p. 40.

32. Wechsler, David. *Manual for the Wechsler Intelligence Scale for Children*, op. cit., p. 15.

33. *Ibid.*, p. 16.

34. Wechsler, David. *The Measurement of Adult Intelligence*, *op*. *cit*., p. 40.

35. Terman, Lewis M. *The Measurement of Intelligence*, Houghton Mifflin Company, New York, 1916, p. 79.

36. Corsini, Raymond. "The Semantics of Intellectual Classification", *Journal of Consulting Psychology*, XV, 1951, p. 487.

37. Krugman, Judith and others. "Pupil Functioning on the Stanford-Binet and the WISC", *Journal of Consulting Psychology*, XV, 1951, p. 481.

38. Weider, A. W. and others. "The Wechsler Intelligence Scale for Children and the Revised Stanford-Binet", *Journal of Consulting Psychology*, XV, 1951, p. 331.

39. Seashore, Harold and others. "The Standardization of the Wechsler Intelligence Scale for Children", *op*. *cit*., p. 109.

40. Garrett, H. E. *Statistics in Psychology and Education*, Longmans, Green and Company, New York, 3rd Ed., 1947, p. 116.

41. *Ibid*., p. 43.

42. *Ibid*., p. 61.

43. Wechsler, David. *Manual for the Wechsler Intelligence Scale for Children*, op. cit., p. 16.

44. Kendall, Maurice G. *Rank Correlation Methods*, Charles Griffin and Company, London, 1948, p. 1.

45. *Ibid*., p. 105.

46. *Ibid*., p. 4.

47. *Ibid*., p. 5.

48. *Ibid*., p. 26.

49. *Ibid*., p. 40.

50. Ibid., p. 39.

51. Ibid., p. 43.

52. Ibid., p. 143.

53. Wechsler, David. op. cit., p. 9.

54. McNemar, Quinn. Psychological Statistics, J. Wiley and Sons, Inc., New York, 1949, p. 139.

55. Wechsler, David. The Measurement of Adult Intelligence, op. cit., p. 224.

56. Yule, G. U. and Kendall, M. G. Introduction to the Theory of Statistics, Charles Griffin and Company, London, 11th Ed., 1937, p. 250.

57. Ibid., p. 251.

58. Ibid.

59. Kendall, M. G. "A New Measure of Rank Correlation", Biometrika, XXX, 1938, p. 85.

60. Kendall, M. G. and others. "The Distribution of Spearman's Coefficient of Rank Correlation in a Universe in Which All Rankings Occur an Equal Number of Times", Biometrika, XXX, 1938, p. 268.

61. Daniels, H. E. "The Relation between Measures of Correlation in the Universe of Sample Permutations", Biometrika, XXXIII, 1944, p. 129.

62. Garrett, H. E. op. cit., p. 198.

63. Ibid., p. 208.

64. Brookes, B. C. and Dick, W. F. L. Introduction to Statistical Method, Parts One and Two, William Heinemann Ltd., London, 1951, p. 254.

65. Ibid., p. 258.

66. Wechsler, David. Manual for the Wechsler Intelligence Scale for Children, op. cit., p. 15.

67. Seashore, Harold G. "Differences between Verbal and Performance IQ's on the Wechsler Intelligence Scale for Children", *op. cit.*, p. 63.

68. Wechsler, David. *op. cit.*, p. 9.

69. Kendall, Maurice G. *Rank Correlation Methods*, *op. cit.*, p. 26.

70. Seashore, Harold and others. "The Standardization of the Wechsler Intelligence Scale for Children", *op. cit.*, p. 103.

71. Garrett, H. E. *op. cit.*, p. 394.

72. Wechsler, David. *op. cit.*, p. 12.

73. Stacey, C. L. and Levin, Janice. *op. cit.*

74. Taylor, L. G. *op. cit.*,

75. Seashore, Harold and others. *op. cit.*, p. 103.

76. Garrett, H. E. *op. cit.*, p. 43.

77. *Ibid.*, p. 61.

78. Wechsler, David. The Measurement of Adult Intelligence, *op. cit.*, p. 6.

79. *Ibid.*

80. Wechsler, David. *Manual for the Wechsler Intelligence Scale for Children*, op. cit., p. 15.

81. *Ibid.*, p. 24.

82. Brookes, B. C. and Dick, W. F. L., *op. cit.*, p. 258.

83. Wechsler, David. *op. cit.*, p. 12.

84. Wechsler, David. *The Measurement of Adult Intelligence*, *op. cit.*, p. 147.

85. Seashore, Harold and others. *op. cit.*, p. 106.

86. Garrett, H. E. *op. cit.*, p. 202.

87. Seashore, Harold G. "Differences between Verbal and Performance IQ's on the Wechsler Intelligence Scale for Children", *op. cit.*, p. 63.

88. *Ibid.*

89. Taylor, L. B. *op. cit.*

90. Seashore, Harold G. *op. cit.*, p. 64.

91. *Ibid.*, p. 63.

92. *Ibid.*, p. 64.

93. Wechsler, David. *op. cit.*, p. 126.

94. Chronbach, Lee J. *op. cit.*, p. 158.

95. Wechsler, David. *op. cit.*, p. 144.

96. *Ibid.*, p. 145.

97. *Ibid.*, p. 144.

98. Kendall, Maurice G. *op. cit.*, p. 25.

BIBLIOGRAPHY

American Association for Gifted Children. The Gifted Child. D. C. Heath and Company, Boston, 1951. 338 pp.

Brookes, B. C. and Dick, W. F. L. Introduction to Statistical Method, Parts One and Two. William Heinemann Ltd., London, 1951. 288 pp.

Chronbach, Lee J. Essentials of Psychological Testing. Harper and Brothers, New York, 1949. 475 pp.

Corsini, Raymond. "The Semantics of Intellectual Classification". Journal of Consulting Psychology XV, 1951. pp. 487-491.

Daniels, H. E. "The Relation between Measures of Correlation in the Universe of Sample Permutations". Biometrika XXXIII, 1944. p. 129.

Garrett, H. E. Statistics in Psychology and Education. Longmans, Green and Company, New York, 3rd Ed., 1947. 487 pp.

Kendall, Maurice G. Rank Correlation Methods. Charles Griffin and Company, London, 1948. 160 pp.

Kendall, M. G. "A New Measure of Rank Correlation". Biometrika XXX, 1938. pp. 81-93.

Kendall, M. G. and others. "The Distribution of Spearman's Coefficient of Rank Correlation in a Universe in Which All Rankings Occur an Equal Number of Times". Biometrika XXX, 1938. pp. 251-273.

Krugman, Judith and others. "Pupil Functioning on the Stanford-Binet and the WISC". Journal of Consulting Psychology XV, 1951. pp. 475-483.

Louttit, C. M. Clinical Psychology. Harper and Brothers, New York and London, Rev. Ed., 1947. 661 pp.

McNemar, Quinn. Psychological Statistics. J. Wiley and Sons, Inc., New York, 1949. 364 pp.

Peterson, Joseph. Early Conceptions and Tests of Intelligence. World Book Company, Yonkers-on-Hudson, New York, and Chicago, 1925. 320 pp.

Richards, T. W. Modern Clinical Psychology. McGraw-Hill Book Company, New York and London, 1946. 331 pp.

Seashore, Harold G. "Differences between Verbal and Performance IQ's on the Wechsler Intelligence Scale for Children". Journal of Consulting Psychology XV, 1951. pp. 62-67.

Seashore, Harold and others. "The Standardization of the Wechsler Intelligence Scale for Children". Journal of Consulting Psychology XIV, 1950. pp. 99-110.

Stacey, C. L. and Levin, Janice. "Correlation Analysis of Scores of Subnormal Subjects on the Stanford-Binet and WISC". American Journal of Mental Deficiency LV, 1951. pp. 590-597.

Taylor, L. B. "An Investigation into the Nature of Results Obtained on the WISC by Mentally Deficient Children". Unfinished master's thesis, (Alberta).

Terman, Lewis M. The Measurement of Intelligence. Houghton Mifflin Company, New York, 1916. 362 pp.

Wechsler, David. The Measurement of Adult Intelligence. The Williams and Wilkins Company, Baltimore, 3rd Ed., 1944. 258 pp.

Wechsler, David. Manual for the Wechsler Intelligence Scale for Children. The Psychological Corporation, New York, 1949. 113 pp.

Weider, A. W. and others. "The Wechsler Intelligence Scale for Children and the Revised Stanford-Binet". Journal of Consulting Psychology XV, 1951. pp. 330-333.

Yule, G. U. and Kendall, M. G. Introduction to the Theory of Statistics. Charles Griffin and Company, London, 11th Ed., 1937. 570 pp.

APPENDIX A

WEIGHTED SCORES AND INTELLIGENCE QUOTIENTS OF THE 57 SUBJECTS ATTAINING 120 I.Q. ON THE WISC FULL SCALE

TABLE XXVI. SEX, CHRONOLOGICAL AGE, AND SUBTEST SCALED SCORES OF 57 SUPERIOR CHILDREN TESTED ON THE WISC

Case Number	Sex	Chronological Age	Subtest Scaled Scores										
			Verbal						Performance				
			Information (1)	Comprehension (2)	Arithmetic (3)	Similarities (4)	Vocabulary (5)	Digit Span (6)	Picture Completion (7)	Picture Arrangement (8)	Block Design (9)	Object Assembly (10)	Coding (11)
1	F	12-10	16	19	16	14	14	15	18	16	16	13	14
2	M	11- 4	16	20	14	15	14	14	16	14	13	11	10
3	M	12- 0	12	18	17	17	13	14	12	13	14	11	10
4	M	12-10	17	16	15	16	14	12	16	16	15	15	14
5	M	12- 2	18	19	15	20	16	14	15	11	12	9	12
6	F	12- 0	13	18	13	16	13	10	16	11	13	12	14
7	F	11-11	15	14	12	15	15	12	13	13	11	11	15
8	M	9- 3	15	16	15	11	16	--	15	15	15	12	11
9	F	13- 1	13	13	16	9	13	12	11	12	18	13	13
10	F	11-10	12	14	14	12	12	16	9	15	13	12	17
11	F	11- 8	13	15	14	13	12	10	16	14	18	12	14
12	F	11- 3	16	17	16	13	13	12	8	10	12	10	12
13	M	11- 3	16	17	18	11	12	14	8	16	17	17	15
14	M	13- 1	13	13	14	15	12	11	14	11	14	13	12
15	M	11- 8	13	16	12	16	13	14	12	14	12	15	12
16	F	10-10	12	12	15	12	14	18	10	14	15	18	12
17	M	11-11	13	16	12	12	13	12	12	14	17	12	13
18	F	9- 6	16	14	13	12	14	18	9	15	11	11	13
19	F	13- 4	14	13	16	14	15	14	15	11	16	13	9
20	F	9- 7	9	17	16	10	15	8	15	13	12	12	15
21	F	9- 7	13	18	14	13	15	13	11	12	14	11	13
22	M	14- 1	16	19	18	16	14	19	8	9	12	9	16
23	F	9- 7	16	16	14	16	15	12	11	17	14	9	10
24	M	9- 7	18	16	18	15	17	13	18	11	17	14	14
25	F	12-11	15	14	16	13	14	14	7	11	10	12	17

TABLE XXVI. (Continued)

Case Number	Sex	Chrono-logical Age	Subtest Scaled Scores										
			Verbal						Performance				
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
26	F	13- 0	16	12	16	16	16	12	14	15	17	12	17
27	M	13- 2	13	17	18	16	14	15	12	13	17	10	14
28	M	13- 3	13	14	10	16	14	10	19	13	14	14	12
29	M	13-11	15	14	14	18	15	10	11	10	17	15	8
30	F	13- 6	14	14	16	14	12	10	15	12	13	12	13
31	M	13- 4	18	13	18	14	17	12	17	12	18	14	12
32	F	12- 1	13	15	12	13	13	8	13	12	15	15	18
33	F	12- 7	13	13	13	11	13	15	15	17	13	15	14
34	M	12- 2	15	17	13	17	15	14	12	11	13	10	11
35	F	12- 6	13	15	15	11	13	12	13	13	10	11	14
36	F	14-10	13	14	15	13	15	19	10	11	13	11	14
37	M	13- 0	15	14	18	12	12	12	11	12	14	12	11
38	F	13- 5	14	19	14	14	15	14	12	15	14	16	13
39	F	14- 8	13	14	13	16	11	14	11	14	12	12	15
40	F	13- 5	12	12	12	16	11	14	12	13	15	15	15
41	M	12- 2	14	15	19	15	12	13	11	11	16	15	14
42	M	12- 6	15	14	16	12	14	17	15	6	16	19	12
43	M	13- 7	17	16	14	17	16	18	9	6	13	12	9
44	M	13-10	16	16	16	17	17	15	17	13	15	16	11
45	F	13- 3	13	13	18	10	14	14	17	13	14	13	11
46	M	14-11	10	12	13	12	10	12	14	16	15	14	20
47	F	13-11	18	16	18	20	18	12	17	15	16	13	14
48	M	7-10	18	17	16	20	17	--	14	11	17	10	8
49	M	11- 7	14	18	16	14	12	10	16	11	12	11	9
50	M	11-10	17	14	17	16	11	14	17	11	16	14	11
51	M	12- 6	10	13	16	13	9	--	15	12	14	17	14
52	F	15- 2	14	19	13	13	13	13	14	14	10	14	10
53	F	11- 2	16	16	10	12	12	14	12	13	12	16	10
54	M	14- 6	18	16	13	18	15	--	12	14	13	13	8
55	M	8- 7	12	13	10	14	15	13	17	12	18	17	18
56	M	14- 4	18	14	18	18	15	--	16	15	18	14	17
57	M	14- 7	13	16	18	17	13	17	11	10	13	14	14

TABLE XXVII. TOTAL SCALED SCORES AND VERBAL, PERFORMANCE, AND FULL SCALE INTELLIGENCE QUOTIENTS OF 57 SUPERIOR CHILDREN TESTED ON THE WISC

	Scaled Score Totals			Intelligence Quotients		
Case Number	Total Verbal (1)	Total Performance (2)	Total Scaled Score (3)	Verbal I.Q. (4)	Performance I.Q. (5)	Full Scale I.Q. (6)
1	79	77	156	137	138	141
2	79	64	143	137	120	131
3	77	60	137	134	114	127
4	78	76	154	135	136	139
5	88	59	147	148	113	134
6	73	66	139	129	122	128
7	71	63	134	126	118	125
8	73	68	141	129	125	130
9	64	67	131	118	124	123
10	64	66	130	118	122	122
11	67	74	141	121	133	130
12	75	52	127	131	103	120
13	74	73	147	130	132	134
14	67	64	131	121	120	123
15	70	65	135	125	121	125
16	65	69	134	119	127	125
17	66	68	134	120	125	125
18	69	59	128	124	113	120
19	72	64	136	128	120	126
20	67	67	134	121	124	125
21	73	61	134	129	114	125
22	83	54	137	142	106	127
23	77	61	138	134	115	128
24	84	74	158	143	133	142
25	72	57	129	128	110	121
26	76	75	151	133	135	137
27	78	66	144	135	122	132
28	67	72	139	121	131	128
29	76	61	137	133	115	127
30	70	65	135	125	121	125
31	80	73	153	138	132	138
32	66	73	139	120	132	128
33	63	74	137	116	133	127
34	77	57	134	134	110	125
35	67	61	128	121	115	120
36	70	59	129	125	113	121
37	71	60	131	126	114	123
38	76	70	146	133	128	133
39	67	64	131	121	120	123
40	63	70	133	116	128	124
41	75	67	142	131	124	131
42	71	68	139	126	125	128
43	80	49	129	138	99	121
44	82	72	154	140	131	139
45	68	69	137	123	127	127
46	57	79	136	109	140	126
47	90	75	165	150	135	147
48	88	60	148	148	114	135
49	74	59	133	130	113	124
50	75	69	144	131	127	132
51	61	72	133	114	131	124
52	72	62	134	128	117	125
53	66	63	129	120	118	121
54	80	60	140	138	114	129
55	64	72	136	118	131	126
56	83	80	163	142	142	146
57	77	62	139	134	117	128

APPENDIX B

RANKS ASSIGNED TO THE SUPERIOR SUBJECTS AS DETERMINED BY WEIGHTED SCORES OBTAINED ON THE SUBTESTS AND ON THE VARIOUS COMBINATIONS OF SUBTESTS

TABLE XXVIII. METHOD OF COMPUTING THE RANKS OF 57 SUBJECTS BASED ON SCALED SCORES OBTAINED ON THE WISC VOCABULARY TEST

Scaled Score Obtained	Number of Subjects Obtaining Score	Cumulative Frequency	Rank Assigned to Subject
18	1	1	1
17	4	5	3.5
16	4	9	7.5
15	12	21	15.5
14	11	32	27
13	11	43	38
12	9	52	48
11	3	55	54
10	1	56	56
9	1	57	57

Kendall dealt with the problem of tied ranks by employing what he called the 'mid-rank method'.

> When there is nothing to choose between individuals, we must clearly rank them all alike if we rank them at all; and our method has the advantage that the sum of the ranks for all members remains the same as for an untied ranking.[98]

In the current situation, 4 subjects obtained weighted or scaled scores of 17 points on the Vocabulary subtest. The rank of 3.5 assigned to each of these subjects was computed by adding the ranks of 2, 3, 4 and 5 which would have been allotted had the scores been distinguishable, and dividing the sum by 4.

Again, 4 subjects obtained scaled scores of 16 on the subtest. Each was assigned the rank $\frac{1}{4}(6 + 7 + 8 + 9) = 7\frac{1}{2}$. By a similar method, tied ranks were computed for subjects who obtained scores of 15, 14, 13, 12 and 11.

TABLE XXIX. RANKS ASSIGNED TO 57 MENTALLY SUPERIOR SUBJECTS ACCORDING TO ATTAINMENT ON EACH OF 10 WISC SUBTESTS, AND ON THE TOTALS OF THE VERBAL AND PERFORMANCE SCALES

Case Number	WISC Subtest										Verbal Scale	Performance Scale
	Verbal					Performance						
	Information (1)	Comprehension (2)	Arithmetic (3)	Similarities (4)	Vocabulary (5)	Picture Completion (6)	Picture Arrangement (7)	Block Design (8)	Object Assembly (9)	Coding (10)		
1	15.5	4	20	31	27	2.5	4.5	15.5	27	19	11.5	3
2	15.5	1	37	25	27	12.5	17.5	39.5	46.5	48	11.5	34.5
3	52	8.5	12.5	9	38	36	26.5	30.5	46.5	48	16.5	46.5
4	9	22.5	29.5	17	27	12.5	4.5	22.5	12	19	13.5	4
5	4	4	29.5	2	7.5	19.5	45.5	48.5	56	35.5	2.5	50.5
6	41.5	8.5	45.5	17	38	12.5	45.5	39.5	36.5	19	28	29
7	24	39	52	25	15.5	30	26.5	53.5	46.5	10	34	37.5
8	24	22.5	29.5	52.5	7.5	19.5	10	22.5	36.5	42.5	28	23
9	41.5	49.5	20	57	38	44.5	35.5	3	27	28.5	52	26
10	52	39	37	46.5	48	52	10	39.5	36.5	4.5	52	29
11	41.5	30.5	37	38.5	48	12.5	17.5	3	36.5	19	43.5	8
12	15.5	13.5	20	38.5	38	55	53	48.5	52.5	35.5	23	56
13	15.5	13.5	6.5	52.5	48	55	4.5	9	4	10	25.5	11

TABLE XXIX. (Continued)

Case Number	WISC Subtest										Verbal Scale	Performance Scale
	Verbal					Performance						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		
14	41.5	49.5	37	25	48	26	45.5	30.5	27	35.5	43.5	34.5
15	41.5	22.5	52	17	38	36	17.5	48.5	12	35.5	37	31.5
16	52	55.5	29.5	46.5	27	49.5	17.5	22.5	2	35.5	50	20
17	41.5	22.5	52	46.5	38	36	17.5	9	36.5	28.5	48	23
18	15.5	39	45.5	46.5	27	52	10	53.5	46.5	28.5	39	50.5
19	30.5	49.5	20	31	15.5	19.5	45.5	15.5	27	52	31	34.5
20	57	13.5	20	55.5	15.5	19.5	26.5	48.5	36.5	10	43.5	26
21	41.5	8.5	37	38.5	15.5	44.5	35.5	30.5	46.5	28.5	28	42.5
22	15.5	4	6.5	17	27	55	55	48.5	56	7	5.5	55
23	15.5	22.5	37	17	15.5	44.5	1.5	30.5	56	48	16.5	42.5
24	4	22.5	6.5	25	3.5	2.5	45.5	9	19.5	19	4	8
25	24	39	20	38.5	27	57	45.5	56	36.5	4.5	31	53.5
26	15.5	55.5	20	17	7.5	26	10	9	36.5	4.5	20	5.5
27	41.5	13.5	6.5	17	27	36	26.5	9	52.5	19	13.5	29
28	41.5	39	56	17	27	1	26.5	30.5	19.5	35.5	43.5	14.5
29	24	39	37	5	15.5	44.5	53	9	12	55.5	20	42.5
30	30.5	39	20	31	48	19.5	35.5	39.5	36.5	28.5	37	31.5
31	4	49.5	6.5	31	3.5	6.5	35.5	3	19.5	35.5	9	11
32	41.5	30.5	52	38.5	38	30	35.5	22.5	12	2	48	11
33	41.5	49.5	45.5	52.5	38	19.5	1.5	39.5	12	19	54.5	8
34	24	13.5	45.5	9	15.5	36	45.5	39.5	52.5	42.5	16.5	53.5
35	41.5	30.5	29.5	52.5	38	30	26.5	56	46.5	19	43.5	42.5

TABLE XXIX. (Continued)

Case Number	WISC Subtest										Verbal Scale	Performance Scale
	Verbal					Performance						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		
36	41.5	39	29.5	38.5	15.5	49.5	45.5	39.5	46.5	19	37	50.5
37	24	39	6.5	46.5	48	44.5	35.5	30.5	36.5	42.5	34	46.5
38	30.5	4	37	31	15.5	36	10	30.5	7	28.5	20	17.5
39	41.5	39	45.5	17	54	44.5	17.5	48.5	36.5	10	43.5	34.5
40	52	55.5	52	17	54	36	26.5	22.5	12	10	54.5	17.5
41	30.5	30.5	1	25	48	44.5	45.5	15.5	12	19	23	26
42	24	39	20	46.5	27	19.5	56.5	15.5	1	35.5	34	23
43	9	22.5	37	9	7.5	52	56.5	39.5	36.5	52	9	57
44	15.5	22.5	20	9	3.5	6.5	26.5	22.5	7	42.5	7	14.5
45	41.5	49.5	6.5	55.5	27	6.5	26.5	22.5	27	42.5	40	20
46	55.5	55.5	45.5	46.5	56	26	4.5	22.5	19.5	1	57	2
47	4	22.5	6.5	2	1	6.5	10	15.5	27	19	1	5.5
48	4	13.5	20	2	3.5	26	45.5	9	52.5	55.5	2.5	46.5
49	30.5	8.5	20	31	48	12.5	45.5	48.5	46.5	52	25.5	50.5
50	9	39	12.5	17	54	6.5	45.5	15.5	19.5	42.5	23	20
51	55.5	49.5	20	38.5	57	19.5	35.5	30.5	4	19	56	14.5
52	30.5	4	45.5	38.5	38	26	17.5	56	19.5	48	31	39.5
53	15.5	22.5	56	46.5	48	36	26.5	48.5	7	48	48	37.5
54	4	22.5	45.5	5	15.5	36	17.5	39.5	27	55.5	9	46.5
55	52	49.5	56	31	15.5	6.5	35.5	3	4	55.5	52	14.5
56	4	39	6.5	5	15.5	12.5	10	3	19.5	4.5	5.5	1
57	41.5	22.5	6.5	9	38	44.5	53	39.5	19.5	19	16.5	39.5

TABLE XXX. RANKS ASSIGNED TO 57 SUPERIOR SUBJECTS ACCORDING TO ACHIEVEMENT ON VARIOUS COMBINATIONS OF 9 WISC SUBTESTS

Case Number	Rank Based on the Sum of Ten Scaled Scores Minus the Scaled Score on:									
	Information (1)	Comprehension (2)	Arithmetic (3)	Similarities (4)	Vocabulary (5)	Picture Completion (6)	Picture Arrangement (7)	Block Design (8)	Object Assembly (9)	Coding (10)
1	3.5	8	3.5	4	3	5.5	6	4	4	5
2	16.5	27	13	14	16.5	19	17.5	12.5	12	12
3	24.5	37.5	37	39.5	27	24.5	29	29	23	24
4	6	6.5	5	6	5	5.5	7	5.5	6	7.5
5	13	12	10	18.5	12	12.5	9.5	7.5	9	9
6	20.5	33	20.5	25.5	19.5	29	19.5	18.5	21	28.5
7	44	35.5	30.5	43	45.5	37.5	38	29	30.5	45
8	20.5	18	20.5	11	23.5	22	25	18.5	17	16.5
9	47.5	40.5	51	30.5	49.5	43.5	48.5	57	47	48.5
10	47.5	47.5	49	46	49.5	37.5	55.5	50.5	47	56
11	14.5	16	16.5	14	16.5	24.5	22	29	17	24
12	57	57	57	56	55	50	52.5	56	52	53
13	10.5	10.5	13	8	8.5	4	14.5	12.5	14	14.5
14	47.5	40.5	47	52.5	45.5	55	44	50.5	47	45
15	33.5	37.5	27	43	34	29	38	29	40	33
16	33.5	30.5	41	30.5	41.5	26.5	44	42.5	55.5	35.5
17	39	40.5	30.5	30.5	38	33	44	50.5	35.5	38.5
18	55.5	53	51	52.5	55	50	57	50.5	52	53
19	33.5	27	37	30.5	38	37.5	27	39	30.5	24
20	24.5	44.5	44.5	23.5	45.5	50	38	33	35.5	45
21	39	47.5	37	36.5	45.5	29	34.5	39	30.5	38.5
22	39	40.5	41	36.5	30.5	15	19.5	21.5	19	38.5
23	33.5	30.5	23.5	30.5	30.5	19	38	24.5	17	20
24	3.5	3	3.5	3	4	3	3	3	3	3

TABLE XXX. (Continued)

Case Number	Rank Based on the Sum of Ten Scaled Scores Minus the Scaled Score on:									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
25	53	50.5	55	52.5	52.5	33	50.5	42.5	52	57
26	7.5	5	7.5	9	8.5	7.5	9.5	9	6	10
27	10.5	14	20.5	14	14.5	12.5	14.5	15.5	11	16.5
28	20.5	18	13	25.5	23.5	43.5	25	21.5	24.5	24
29	33.5	27	27	43	34	22	22	39	35.5	18
30	39	33	41	36.5	30.5	43.5	32	33	30.5	35.5
31	7.5	4	7.5	5	7	9	4.5	7.5	6	6
32	20.5	22	16.5	21.5	19.5	22	22	24.5	27	38.5
33	27.5	22	23.5	21.5	27	33	44	24.5	35.5	33
34	44	44.5	33.5	48.5	45.5	33	32	36	27	33
35	52	55	55	48.5	52.5	57	55.5	46	52	55
36	50.5	50.5	53	52.5	55	50	50.5	54.5	47	53
37	50.5	44.5	43	43	45.5	43.5	48.5	50.5	43	41.5
38	9	14	10	10	12	10.5	14.5	10	14	12
39	47.5	44.5	44.5	55	41.5	43.5	52.5	42.5	43	50.5
40	39	33	33.5	48.5	34	37.5	44	46	47	48.5
41	14.5	14	27	18.5	14.5	14	14.5	18.5	21	20
42	27.5	18	27	18.5	23.5	26.5	11.5	29	40	24
43	55.5	55	51	57	57	43.5	32	54.5	52	41.5
44	5	6.5	6	7	6	7.5	4.5	5.5	9	4
45	27.5	22	41	18.5	30.5	43.5	29	33	27	27
46	20.5	22	27	23.5	19.5	33	44	36	35.5	50.5
47	1	1.5	1	1.5	2	1	1	1	1	1
48	12	9	10	14	12	10.5	8	11	9	7.5
49	44	50.5	47	43	38	55	34.5	36	35.5	30.5
50	16.5	10.5	16.5	14	10	19	11.5	14	14	12
51	30	35.5	47	39.5	27	53	38	42.5	55.5	45
52	42	50.5	33.5	36.5	38	43.5	44	24.5	40	30.5
53	54	55	41	48.5	51	55	54	50.5	57	45
54	33.5	22	16.5	30.5	23.5	16.5	25	15.5	21	14.5
55	27.5	27	20.5	30.5	38	50	29	46	43	20
56	2	1.5	2	1.5	1	2	2	2	2	2
57	20.5	27	33.5	30.5	19.5	16.5	17.5	18.5	24.5	28.5

TABLE XXXI. RANKS ASSIGNED TO 57 SUPERIOR SUBJECTS ACCORDING TO ACHIEVEMENT ON VARIOUS COMBINATIONS OF 4 WISC VERBAL SUBTESTS

Case Number	Rank Based on the Sum of Five Verbal Scaled Scores Minus the Scaled Score on:				
	Information	Comprehension	Arithmetic	Similarities	Vocabulary
1	12	19	13	8	8.5
2	12	22	9	10	8.5
3	8	22	19.5	23	13.5
4	19.5	12.5	13	15.5	13.5
5	2.5	3.5	1	3.5	1.5
6	23.5	36.5	19.5	36.5	26.5
7	37.5	28.5	24	40	37.5
8	29.5	28.5	28.5	15.5	34.5
9	53	48.5	55	42	53.5
10	50.5	53	52.5	50	50.5
11	43	45.5	46	44.5	40.5
12	26	24.5	24	15.5	21
13	29.5	28.5	34.5	11.5	21
14	43	39.5	46	50	40.5
15	34	39.5	28.5	44.5	34.5
16	47.5	42.5	52.5	47.5	53.5
17	47.5	53	42	44.5	47
18	47.5	36.5	34.5	36.5	40.5
19	29.5	22	34.5	33	34.5
20	29.5	53	49.5	36.5	50.5
21	23.5	36.5	24	23	31
22	4	9.5	9	5	4
23	19.5	15.5	13	19	21
24	5.5	5	6	2	6
25	34	24.5	34.5	29	31
26	23.5	9.5	19.5	23	26.5
27	8	15.5	19.5	15.5	13.5
28	43	42.5	31	52.5	47

TABLE XXXI. (Continued)

Case Number	Rank Based on the Sum of Five Verbal Scaled Scores Minus the Scaled Score on:				
	Information	Comprehension	Arithmetic	Similarities	Vocabulary
29	29.5	12.5	16	33	24.5
30	37.5	33	42	40	31
31	15.5	6	16	6	17.5
32	47.5	48.5	42	47.5	47
33	55.5	53	52.5	50	55
34	15.5	19	11	23	21
35	43	45.5	48	40	44
36	34	33	38.5	36	40.5
37	37.5	28.5	46	29	28
38	15.5	28.5	16	15.5	24.5
39	43	42.5	42	52.5	37.5
40	53	48.5	49.5	56	50.5
41	19.5	19	34.5	23	17.5
42	37.5	28.5	38.5	29	34.5
43	12	9.5	6	11.5	13.5
44	5.5	7	6	8	8.5
45	40	36.5	52.5	33	44
46	57	57	57	57	57
47	1	1	2.5	1	1.5
48	2.5	2	2.5	3.5	3
49	23.5	33	28.5	23	21
50	29.5	15.5	28.5	29	13.5
51	53	56	56	55	50.5
52	29.5	42.5	24	29	28.5
53	55.5	53	34.5	44.5	44
54	15.5	9.5	4	15.5	8.5
55	50.5	48.5	42	54	56
56	8	3.5	9	8	5
57	10	15.5	24	23	13.5

TABLE XXXII. RANKS ASSIGNED TO 57 SUPERIOR SUBJECTS ACCORDING TO ACHIEVEMENT ON VARIOUS COMBINATIONS OF 4 WISC PERFORMANCE SUBTESTS

Case Number	Rank Based on the Sum of Five Performance Scaled Scores Minus the Scaled Score on:				
	Picture Completion	Picture Arrangement	Block Design	Object Assembly	Coding
1	8	6	4	3	2.5
2	48	38.5	34	28.5	26.5
3	48	49	50	42.5	41.5
4	5.5	10.5	4	7	4
5	54.5	45.5	45.5	39	49.5
6	38	23.5	24	25.5	35.5
7	38	38.5	29	31.5	47
8	26	30	24	15.5	19
9	18	23.5	39	25.5	26.5
10	14.5	35	24	25.5	44
11	11.5	10.5	14	5.5	9.5
12	54.5	57	56	56	55
13	1.5	18	14	15.5	15
14	38	30	37	35	25.5
15	26	35	24	39	29.5
16	8	23.5	20	35	19
17	18	26.5	34	15.5	23.5
18	38	54.5	41.5	46.5	51.5
19	44	30	41.5	35	23.5
20	30	26.5	17	21	35.5
21	38	41.5	45.5	39	47
22	51.5	53	55	54.5	57
23	38	54.5	45.5	31.5	40
24	18	2.5	11.5	8	9.5
25	38	51	45.5	54.5	55
26	4	10.5	8.5	4	15
27	23	30	39	15.5	35.5

TABLE XXXII. (Continued)

Case Number	Rank Based on the Sum of Five Performance Scaled Scores Minus the Scaled Score on: Picture Completion	Picture Arrangement	Block Design	Object Assembly	Coding
28	26	14.5	8.5	11.5	9.5
29	38	35	52.5	53	29.5
30	38	30	29	28.5	35.5
31	18	6	17	9.5	6
32	5.5	6	8.5	11.5	23.5
33	8	18	4	9.5	9.5
34	53	51	52.5	51	51.5
35	48	45.5	34	39	49.5
36	44	45.5	50	46.5	53
37	44	45.5	50	46.5	44
38	11.5	23.5	14	25.5	19
39	26	38.5	29	31.5	44
40	11.5	18	17	21	23.5
41	18	20.5	34	31.5	29.5
42	26	4	29	42.5	21
43	57	56	57	57	55
44	21.5	14.5	11.5	15.5	6
45	30	20.5	20	15.5	15
46	1.5	2.5	1	2	12
47	11.5	10.5	6	5.5	6
48	51.5	41.5	54	39	35.5
49	56	45.5	45.5	46.5	41.5
50	30	16	24	21	15
51	14.5	10.5	8.5	21	15
52	48	45.5	29	46.5	35.5
53	32.5	38.5	34	51	29.5
54	48	51	45.5	51	35.5
55	21.5	10.5	20	21	1
56	3	1	2	1	2.5
57	32.5	33	39	46.5	47

TABLE XXXIII. RANKS ASSIGNED TO 52 SUPERIOR SUBJECTS ACCORDING TO SCALED SCORES ATTAINED ON DIGIT SPAN, AND ON 5 VERBAL AND 10 WISC SUBTESTS OMITTING DIGIT SPAN

Case Number	Rank of Scaled Score on Digit Span	5 Verbal Subtests	10 WISC Subtests
1	10.5	8.5	3
2	19.5	8.5	13
3	19.5	13.5	24
4	37.5	10.5	4.5
5	19.5	2	8.5
6	47.5	24.5	18
7	37.5	30	35
8	-	-	-
9	37.5	48	42.5
10	8	48	45
11	47.5	39.5	15
12	37.5	20	52
13	19.5	22.5	8.5
14	44	39.5	42.5
15	19.5	33	30.5
16	4	46	35
17	37.5	44	35
18	4	35	50.5
19	19.5	27	28
20	51.5	39.5	35
21	29	24.5	35
22	1.5	4	24
23	37.5	13.5	21
24	29	3	2
25	19.5	27	47.5
26	37.5	17	7
27	10.5	10.5	11.5
28	47.5	39.5	18
29	47.5	17	24
30	47.5	33	30.5
31	37.5	6.5	6
32	51.5	44	18
33	10.5	50.5	24
34	19.5	13.5	35
35	37.5	39.5	50.5
36	1.5	33	47.5
37	37.5	30	42.5
38	19.5	17	10
39	19.5	39.5	42.5
40	19.5	50.5	39.5
41	29	20	14
42	6.5	30	18
43	4	6.5	47.5
44	10.5	5	4.5
45	19.5	36	24
46	37.5	52	28
47	37.5	1	1
48	-	-	-
49	47.5	22.5	39.5
50	19.5	20	11.5
51	-	-	-
52	29	27	35
53	19.5	44	47.5
54	-	-	-
55	29	48	28
56	-	-	-
57	6.5	13.5	18

APPENDIX C

ILLUSTRATIONS OF STATISTICAL METHODS APPLIED

TABLE XXXIV. TO CALCULATE THE QUANTITY 'S' EMPLOYED IN THE DETERMINATION OF THE COEFFICIENT OF RANK CORRELATION BETWEEN RANKS OBTAINED ON VOCABULARY AND ON THE FOUR REMAINING VERBAL SUBTESTS COMBINED

Case Number (1)	Rank on Vocabulary Subtest (2)	Rank on Other Four Verbal Subtests Combined (3)	Contribution to 'S' (4)
47	1	1.5	+55
24	3.5	6	+46
31	3.5	17.5	+27
44	3.5	8.5	+43
48	3.5	3	+50
5	7.5	1.5	+48
8	7.5	34.5	-3
26	7.5	26.5	+11
43	7.5	13.5	+33
7	15.5	37.5	-3
19	15.5	34.5	0
20	15.5	50.5	-25
21	15.5	31	+4
23	15.5	21	+15
29	15.5	24.5	+12
34	15.5	21	+15
36	15.5	40.5	-7
38	15.5	24.5	+12
54	15.5	8.5	+32
55	15.5	56	-34
56	15.5	5	+34
1	27	8.5	+25
2	27	8.5	+25
4	27	13.5	+22
16	27	53.5	-20
18	27	40.5	-3
22	27	4	+25
25	27	31	+4
27	27	13.5	+22
28	27	47	-11
42	27	34.5	+2
45	27	44	-7
3	38	13.5	+13
6	38	26.5	+6
9	38	53.5	-12
12	38	21	+8
15	38	34.5	+2
17	38	47	-6
32	38	47	-6
33	38	55	-12
35	38	44	-5
52	38	28.5	+5
57	38	13.5	+13
10	48	50.5	-1
11	48	40.5	+1
13	48	21	+3
14	48	40.5	+1
30	48	31	+3
37	48	28.5	+3
41	48	17.5	+3
49	48	21	+3
53	48	44	+1
39	54	37.5	+2
40	54	50.5	+1
50	54	13.5	+2
46	56	57	-1
51	57	50.5	0

Adding column (4), S = +476

STEPS IN THE CALCULATION OF 'S' FOR VOCABULARY VS. SUM OF FOUR REMAINING VERBAL SCALED SCORES

1. Ranks obtained on the Vocabulary subtest (see column headed 'Vocabulary', Table XXIX) were arranged in order of magnitude, beginning with the uppermost rank of one, obtained by subject 47, and ending with the lowest rank of 57, obtained by subject 51. The rearranged Vocabulary ranking is to be found in column (2), Table XXXIV.

2. The rank of each subject based on the sum of his scaled scores on the four remaining Verbal subtests (see appropriate column, Table XXXI) was then paired with his rank obtained on Vocabulary. This ranking appears in column (3), Table XXXIV.

3. Subject 47's contribution to 'S' was now determined. This involved counting the number of times his rank of 1.5 in column (3) exceeded the ranks of other subjects below him in the same column. It may be seen that he was tied in rank only by subject 5, while his rank of 1.5 exceeded those of 55 other subjects. His contribution to 'S', accordingly, was set at +55. This value appears as the first entry in column (4), Table XXXIV.

4. Subjects 24, 31, 44 and 48 were each allotted a rank of 3.5 on Vocabulary. As a result, the contribution to 'S' by each of these subjects is computed only from ranks which fall below the dotted line located under the entries for subject 48.

5. Subject 24 obtained the rank of 6 on the remaining Verbal subtests. This rank exceeds those of 49 of the 52 subjects below the dotted line following subject 48's entries. Only subjects 5, 56 and 22, with ranks of 1.5, 5 and 4 respectively, exceeded subject 24 on the combined tests. Accordingly, subject 24's contribution to 'S' was found to be (49-3) or +46. This value constitutes the second entry in column (4).

6. The remaining 55 entries in column (4) were determined in the manner just described. A negative contribution to 'S' resulted where a subject's rank in column (3) exceeded the ranks below it on fewer occasions than these ranks in turn exceeded the rank of the subject.

7. 'S' was computed by adding the values of the 57 entries in column (4).

COMPUTATION OF THE RANK CORRELATION COEFFICIENT 'TAU' FOR VOCABULARY VS. FOUR REMAINING VERBAL SUBTESTS

Where tied ranks exist in both rankings, Kendall defines his coefficient of rank correlation as

$$\text{tau} = \frac{S}{\sqrt{\frac{1}{2}n(n-1) - T}\ \sqrt{\frac{1}{2}n(n-1) - U}}$$

In the present instance, S = +476 (calculated in Table XXXIV), $\frac{1}{2}n(n-1) = \frac{1}{2}(57)(56) = 1596$, T = 227, and U = 65. T and U are correction terms whose magnitudes are determined by the number of ties in the two rankings.

In the ranking of Vocabulary scores presented in column (2) of Table XXXIV, there are seen to be seven sets of ties: one set composed of three ties, two sets of four apiece, one of nine, two of eleven, and one of twelve.

The correction for ties in the first of a pair of rankings is given by Kendall as

$$T = \frac{1}{2} \Sigma_t t(t-1)$$

Here, the values of 't' to be substituted in the formula were in turn 3, 4, 4, 9, 11, 11, and 12. T then became $\frac{1}{2}[(3)(2) + (4)(3) + (4)(3) + (9)(8) + (11)(10) + (11)(10) + (12)(11)] = 227$.

U was obtained by applying the correction formula for ties in the second of a pair of rankings being correlated. This formula is

$$U = \frac{1}{2} \Sigma_u (u-1)$$

In the second ranking being considered (column 3, Table XXXIV), there are seven sets of two ties, three sets of three ties, four sets of four, one of five, and one of six. The values of 'u' to be substituted in the formula were then 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, and 6. Hence U became $\frac{1}{2}[(7)(2)(1) + (3)(3)(2) + (4)(4)(3) + (1)(5)(4) + (1)(6)(5)] = 65$.

Tau was found by substituting the calculated values of S, $\frac{1}{2}n(n-1)$, T, and U in the formula for rank correlation. Accordingly,

$$\text{tau} = \frac{+476}{\sqrt{1596-227}\ \sqrt{1596-65}} = \frac{+476}{1448} = +0.329$$

TESTING THE SIGNIFICANCE OF THE OBSERVED RANK CORRELATION COEFFICIENT 'TAU' FOR VOCABULARY VS. FOUR REMAINING VERBAL SUBTESTS

The question arose as to whether or not the obtained tau value of +0.329 for Vocabulary vs. the remaining four Verbal subtests indicated that a correlation actually exists in the population from which the sample of 57 was chosen. According to Kendall, a test of tau is equivalent to a test of the corresponding value of S, in this case, +476.

The variance of S was computed by applying the formula

$$\text{Var } S = 1/18\left[n(n-1)(2n+5) - \sum_t t(t-1)(2t+5) - \sum_u u(u-1)(2u+5)\right] + \frac{1}{9n(n-1)(n-2)}\left[\sum_t t(t-1)(t-2)\right]\left[\sum_u u(u-1)(u-2)\right] + \frac{1}{2n(n-1)}\left[\sum_t t(t-1)\right]\left[\sum_u u(u-1)\right]$$

In the present instance, there were ties of extent $t = 3, 4, 4, 9, 11, 11$, and 12 in the first ranking, and of extent $u = 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5$, and 6 in the second ranking.

As a result of these ties,

$$\sum_t t(t-1)(2t+5) = 11802$$

$$\sum_u u(u-1)(2u+5) = 1758$$

$$\sum_t t(t-1)(t-2) = 3858$$

$$\sum_u u(u-1)(u-2) = 294$$

$$\sum_t t(t-1) = 454$$

$$\sum_u u(u-1) = 130$$

Since the value of n was 57

$$n(n-1)(2n+5) = 379848$$

$$9n(n-1)(n-2) = 1580040$$

$$2n(n-1) = 6384$$

On substituting these computed values, the variance of S reduced to

$$\text{Var } S = 1/18\left[379848 - (11802 + 1758)\right] + \frac{(3858)(294)}{1580040} + \frac{(454)(130)}{6384}$$

$$\text{Var } S = 20359.2$$

$$\sigma_s = \sqrt{20359.2} = 142.7$$

With a correction for continuity, $\frac{S - 1}{\sigma_s} = \frac{475}{142.7} = 3.33$

Therefore, S (corrected) = 3.33 σ_s

From Kendall's Appendix Table 3 it may be seen that the probability of a deviation less than 3.33 is about 0.99957. The probability that 3.33 is obtained or exceeded in absolute value is thus about 2(1 - 0.99957) = 0.0008. This is less than 0.001, and it may be concluded that the observed value of S, and consequently that of Tau, is highly significant, and is not different from zero due to accidents of sampling.

TABLE XXXV. TO ILLUSTRATE THE CALCULATION OF THE MEAN AND STANDARD DEVIATION OF 57 SCALED SCORES OBTAINED ON THE VOCABULARY SUBTEST

Scaled Score	f	x^1	fx^1	$f(x^1)^2$
18	1	4	4	16
17	4	3	12	36
16	4	2	8	16
15	12	1	12	12
14	11	0	0	0
13	11	-1	-11	11
12	9	-2	-18	36
11	3	-3	-9	27
10	1	-4	-4	16
9	1	-5	-5	25
Σ's	57		-11	195

$$\Sigma f = N = 57$$

$$\Sigma fx^1 = -11$$

$$\Sigma f(x^1)^2 = 195$$

$$\text{Mean} = \text{Assumed Mean} - (\Sigma fx^1/N)\, i = \text{A.M.} + ci$$

A.M. = 14.00

(ci) = -0.19

$$c = \Sigma fx^1/N = -11/57 = -0.19$$

$$c^2 = (-0.19)^2 = 0.036$$

$$\boxed{\text{Mean} = 13.81}$$

$$\text{S.D.} = \sqrt{\frac{\Sigma(fx^1)^2}{N} - c^2} \times i \text{ (interval)}$$

$$= \sqrt{\frac{195}{57} - 0.036} \times 1$$

$$\boxed{\text{S.D.} = 1.84}$$

CORRELATION CHART FOR COMPUTATION OF PEARSON PRODUCT-MOMENT COEFFICIENT OF CORRELATION

X SCALE REPRESENTS TOTAL PERFORMANCE SCALED SCORES

Y SCALE REPRESENTS TOTAL VERBAL SCALED SCORES

Y						49-51	52-54	55-57	58-60	61-63	64-66	67-69	70-72	73-75	76-78	79-81						f	d	y'	y'^2	$+x'y'$	$-x'y'$
	-100	-90	-80	-70	-60	-50	-40	-30	-20	-10		10	20	30	40	50	60	70	80	90	100		10				
	-90	-81	-72	-63	-54	-45	-36	-27	-18	-9		9	18	27	36	45	54	63	72	81	90		9				
	-80	-72	-64	-56	-48	-40	-32	-24	-16	-8		8	16	24	32	40	48	56	64	72	80		8				
	-70	-63	-56	-49	-42	-35	-28	-21	-14	-7		7	14	21	28	35	42	49	56	63	70		7				
90-92	-60	-54	-48	-42	-36	-30	-24	-18	-12	-6		6	12	18 1	24	30	36	42	48	54	60	1	6	6	36	18	
87-89	-50	-45	-40	-35	-30	-25	-20	-15	-10 2	-5		5	10	15	20	25	30	35	40	45	50	2	5	10	50		20
84-86	-40	-36	-32	-28	-24	-20	-16	-12	-8	-4		4	8	12 1	16	20	24	28	32	36	40	1	4	4	16	12	
81-83	-30	-27	-24	-21	-18	-15	-12 1	-9	-6	-3		3	6 1	9	12	15 1	18	21	24	27	30	3	3	9	27	9	
78-80	-20	-18	-16	-14	-12	-10 1	-8	-6	-4 1	-2	2	2	4	6 1	8 2	10	12	14	16	18	20	7	2	14	28	8	
75-77	-10	-9	-8	-7	-6	-5	-4 1	-3 1	-2 1	-1 3		1 2	2 1	3 1	4	5	6	7	8	9	10	10	1	10	10		5
72-74								1	1	2	2	1		1								8					
69-71	10	9	8	7	6	5	4	3	2 3	1 1	2	-1 1	-2	-3	-4	-5	-6	-7	-8	-9	-10	7	-1	-7	7	6	
66-68	20	18	16	14	12	10	8	6	4	2 2	2	-2 3	-4 1	-6 2	-8	-10	-12	-14	-16	-18	-20	10	-2	-20	40		18
63-65	30	27	24	21	18	15	12	9	6	3	1	-3 2	-6 2	-9 1	-12	-15	-18	-21	-24	-27	-30	6	-3	-18	54		27
60-62	40	36	32	28	24	20	16	12	8	4		-4	-8 1	-12	-16	-20	-24	-28	-32	-36	-40	1	-4	-4	16		8
57-59	50	45	40	35	30	25	20	15	10	5		-5	-10	-15	-20	-25 1	-30	-35	-40	-45	-50	1	-5	-5	25		25
	60	54	48	42	36	30	24	18	12	6		-6	-12	-18	-24	-30	-36	-42	-48	-54	-60		-6				
	70	63	56	49	42	35	28	21	14	7		-7	-14	-21	-28	-35	-42	-49	-56	-63	-70		-7				
	80	72	64	56	48	40	32	24	16	8		-8	-16	-24	-32	-40	-48	-56	-64	-72	-80		-8				
	90	81	72	63	54	45	36	27	18	9		-9	-18	-27	-36	-45	-54	-63	-72	-81	-90		-9				
	100	90	80	70	60	50	40	30	20	10		-10	-20	-30	-40	-50	-60	-70	-80	-90	-100		-10				
f						1	2	2	8	8	9	9	6	8	2	2						Σf (N) 57		$\Sigma y'$ -1	$\Sigma y'^2$ 309	$\Sigma +x'y'$ 53	$\Sigma -x'y'$ 103
d	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1		1	2	3	4	5	6	7	8	9	10						
x'						-5	-8	-6	-16	-8	0	9	12	24	8	10						$\Sigma x'$ 20				$\Sigma x'y'$ -50	
x'^2						25	32	18	32	8	0	9	24	72	32	50						$\Sigma x'^2$ 302					

$$\frac{\Sigma x'}{N} = \frac{20}{57} = 0.35$$

$$\left(\frac{\Sigma x'}{N}\right)^2 = (\ 0.35\ )^2 = 0.12$$

$$\frac{\Sigma x'^2}{N} = \frac{302}{57} = 5.30$$

$$\frac{\Sigma y'}{N} = \frac{-1}{57} = -0.02$$

$$\left(\frac{\Sigma y'}{N}\right)^2 = (\ -0.02\ )^2 = 0.00$$

$$\frac{\Sigma y'^2}{N} = \frac{309}{57} = 5.42$$

$$\frac{\Sigma x'y'}{N} = \frac{-50}{57} = -0.88$$

$$\sigma_x = \sqrt{\frac{\Sigma x'^2}{N} - \left(\frac{\Sigma x'}{N}\right)^2} = \sqrt{5.30-0.12}$$

$$= 2.28$$

$$\sigma_y = \sqrt{\frac{\Sigma y'^2}{N} - \left(\frac{\Sigma y'}{N}\right)^2} = \sqrt{5.42-0.00}$$

$$= 2.33$$

$$r = \frac{\frac{\Sigma x'y'}{N} - \left(\frac{\Sigma x'}{N}\right)\left(\frac{\Sigma y'}{N}\right)}{\sigma_x \cdot \sigma_y}$$

$$= \frac{-0.88 - (.35)(-.02)}{(2.28)\ (2.33)}$$

$$= \frac{-0.87}{5.31} = \boxed{-0.16}$$

$$\sigma_r = \frac{1-r^2}{\sqrt{N}} = \frac{1-0.03}{\sqrt{57}}$$

$$= \frac{0.97}{7.55} = 0.13$$

Figure 4.

TABLE XXXVI. TO DETERMINE THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN COMPREHENSION AND VOCABULARY MEAN SCORES

Mean of 57 Comprehension Scores, $\bar{x}$	15.351
Mean of 57 Vocabulary Scores, $\bar{y}$	13.807
Difference Between Means, $\bar{x} - \bar{y}$ or $\bar{d}$	1.544

Subject Number s	Scaled Score on Comprehension x	Scaled Score on Vocabulary y	Difference between Scores d	Deviation from Mean Difference $(d-\bar{d})$	Deviation Squared $(d-\bar{d})^2$
1	19	14	+5	3.456	11.944
2	20	14	+6	4.456	19.856
3	18	13	+5	3.456	11.944
4	16	14	+2	0.456	0.208
5	19	16	+3	1.456	2.120
6	18	13	+5	3.456	11.944
7	14	15	-1	-2.544	6.472
8	16	16	0	-1.544	2.384
9	13	13	0	-1.544	2.384
10	14	12	+2	0.456	0.208
11	15	12	+3	1.456	2.120
12	17	13	+4	2.456	6.032
13	17	12	+5	3.456	11.944
14	13	12	+1	-0.544	0.296
15	16	13	+3	1.456	2.120
16	12	14	-2	-3.544	12.560
17	16	13	+3	1.456	2.120
18	14	14	0	-1.544	2.384
19	13	15	-2	-3.544	12.560
20	17	15	+2	0.456	0.208

s	x	y	d	$(d-\bar{d})$	$(d-\bar{d})^2$
21	18	15	+3	1.456	2.120
22	19	14	+5	3.456	11.944
23	16	15	-1	-0.544	0.296
24	16	17	-1	-2.544	6.472
25	14	14	0	-1.544	2.384
26	12	16	-4	-5.544	30.736
27	17	14	+3	1.456	2.120
28	14	14	0	-1.544	2.384
29	14	15	-1	-2.544	6.472
30	14	12	+2	0.456	0.208
31	13	17	-4	-5.544	30.736
32	15	13	+2	0.456	0.208
33	13	13	0	-1.544	2.384
34	17	15	+2	0.456	0.208
35	15	13	+2	0.456	0.208
36	14	15	-1	-2.544	6.472
37	14	12	+2	0.456	0.208
38	19	15	+4	2.456	6.032
39	14	11	+3	1.456	2.120
40	12	11	+1	-0.544	0.296
41	15	12	+3	1.456	2.120
42	14	14	0	-1.544	2.384
43	16	16	0	-1.544	2.384
44	16	17	-1	-2.544	6.472
45	13	14	-1	-2.544	6.472
46	12	10	+2	0.456	0.208
47	16	18	-2	-3.544	12.560
48	17	17	0	-1.544	2.384
49	18	12	+6	4.456	19.856
50	14	11	+3	1.456	2.120
51	13	9	+4	2.456	6.032
52	19	13	+6	4.456	19.856
53	16	12	+4	2.456	6.032
54	16	15	+1	-0.544	0.296
55	13	15	-2	-3.544	12.560
56	14	15	-1	-2.544	6.472
57	16	13	+3	1.456	2.120

$$\Sigma(d-\bar{d})^2 = 346.144$$

$$S.E._{\bar{d}} = \sqrt{\frac{\Sigma(d-\bar{d})^2}{n(n-1)}}$$

$$= \sqrt{\frac{346.144}{(57)(56)}}$$

$$= 0.329$$

$$t = \frac{\bar{d}}{S.E._{\bar{d}}} = \frac{1.544}{0.329} = 4.69$$

From a table of t for 56 degrees of freedom, one finds the entries 2.39 and 2.67 at the .02 and .01 levels respectively. Hence the mean difference of +1.544 is significant at the .005 level, and there is reason to believe that superior subjects find the Comprehension subtest less difficult than the Vocabulary subtest.